Contents

Project Summary

Initially, the hope of this project was to be able to compare NAIP imagery from two years and view the development changes, including those of landscaping and swimming pools, and see how much Chico, CA has changed over the past few years, and to see if the effects of the drought were visible in the landscaping efforts. What this ended up being was a crash-course in performing image classifications, and learning about all the things NOT to do when attempting to classify an image. Everything from using the wrong file type, to adding too many classification categories, I tested all the limits. I won't say that this project didn't push me to my limits, too, but it made it that much more exciting when I did finally get to get things to work. And with that, I bring you my image classification… from the high hopes of doing a temporal comparison, to explaining what not to do in the event of performing a classification. Enjoy!

Purpose.

In doing this supervised image classification, it is my hope to see if there is not only an increase in population of Chico, but if it is concentrated in certain areas, or just a general spread of development. Considering the changes in climate the area has seen (wetter years in 2006 and 2011, drought years in 2014 and 2015) (Butte County Department of Water and Resource Conservation, 2016), I'm also looking to see if I can determine if the number of swimming pools and areas with green landscaping within residential areas maintains a constant, or if they change over time. I've chosen to look at NAIP imagery for this process so that I can get the level of detail I'm looking for. The diagram below is from the 2016 Butte County Water Inventory and Analysis Report, and shows the annual water index for the Sacramento River Valley. Note the years of critical water levels in 2014 and 2015, and the decreased number of wet years from 2007 forward.

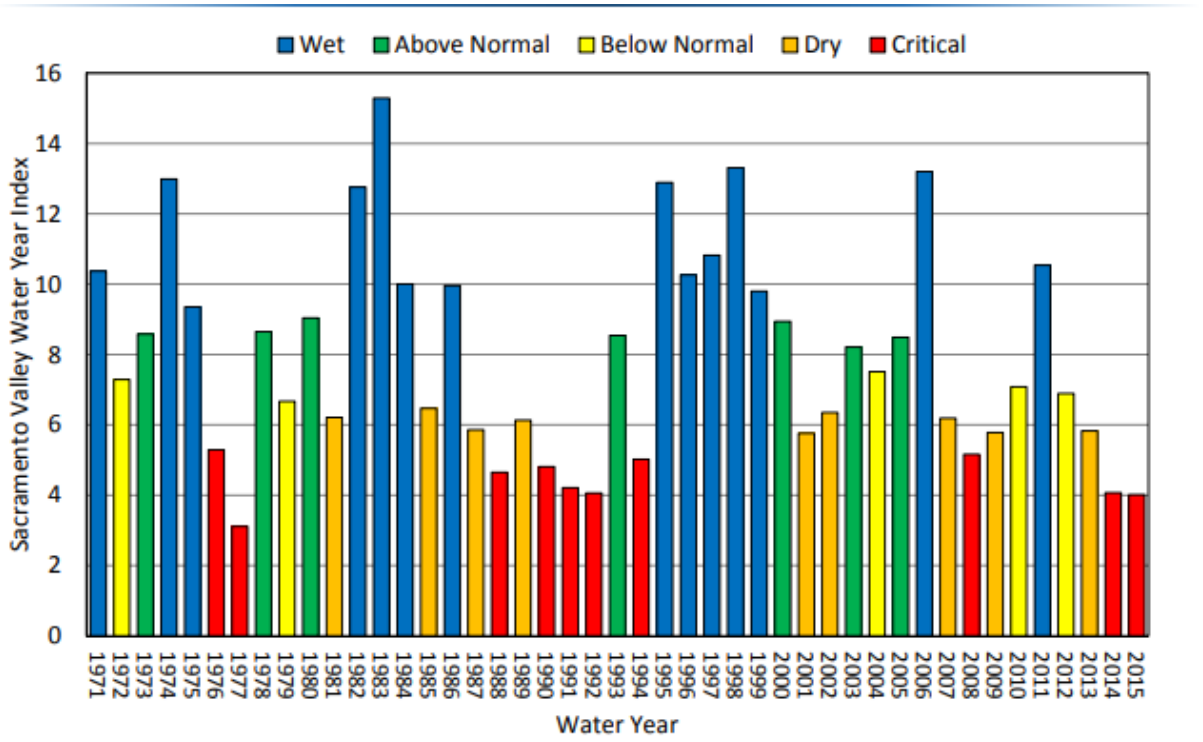


Figure 1. This image from the Butte County Water Inventory and Analysis Report (2016) shows the annual year index for the Sacramento Valley area. To obtain these values, runoff values from the current and previous year are weighted and compared.

Methods

2005 Imagery

Data collection for this imagery was done through the state's web resources, which I will discuss the consequences of this later in this paper. Since this was such a large dataset to work with, it was beneficial to perform clips to the layers so that I was not forcing my computer to process the imagery for the entire state of California while working on this project. To clip the layers, I accessed the "Raster Functions" by clipping on the "Imagery" tab and clicking on "Raster Functions," which opened up a tab on the side where I could search for "Clip." This creates a temporary image that can be used for running processes on a smaller portion of the original image, without creating a permanent copy on your disk.

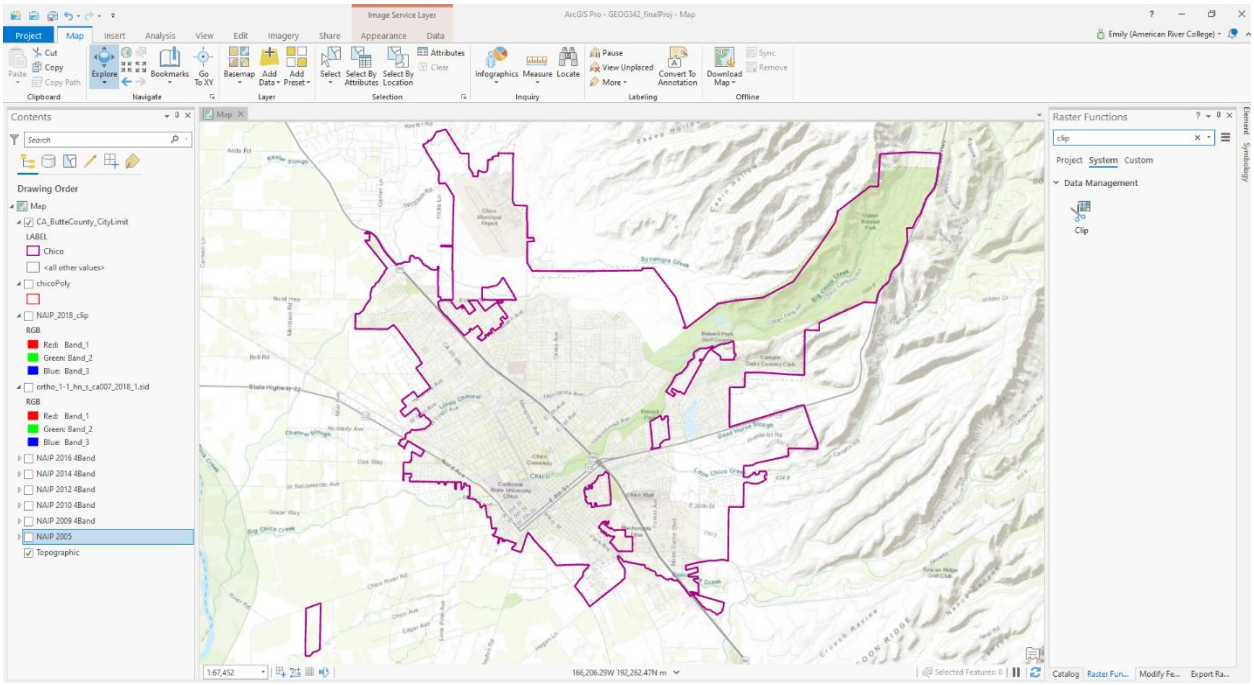


Figure 2. Once the desired layers are added in ArcPro, they can be clipped to a desired size which helps to speed up the map redraw.

The use of the clip tool is fairly self-explanatory. Either you're looking to clip portions that fall inside of your boundary (this can be a screen setting, or a feature added in ArcPro), or outside. In this case, I was looking to have the features I was viewing on the screen, so I chose "Outside" for the clip type, since I was looking to remove those areas outside of my field of view.

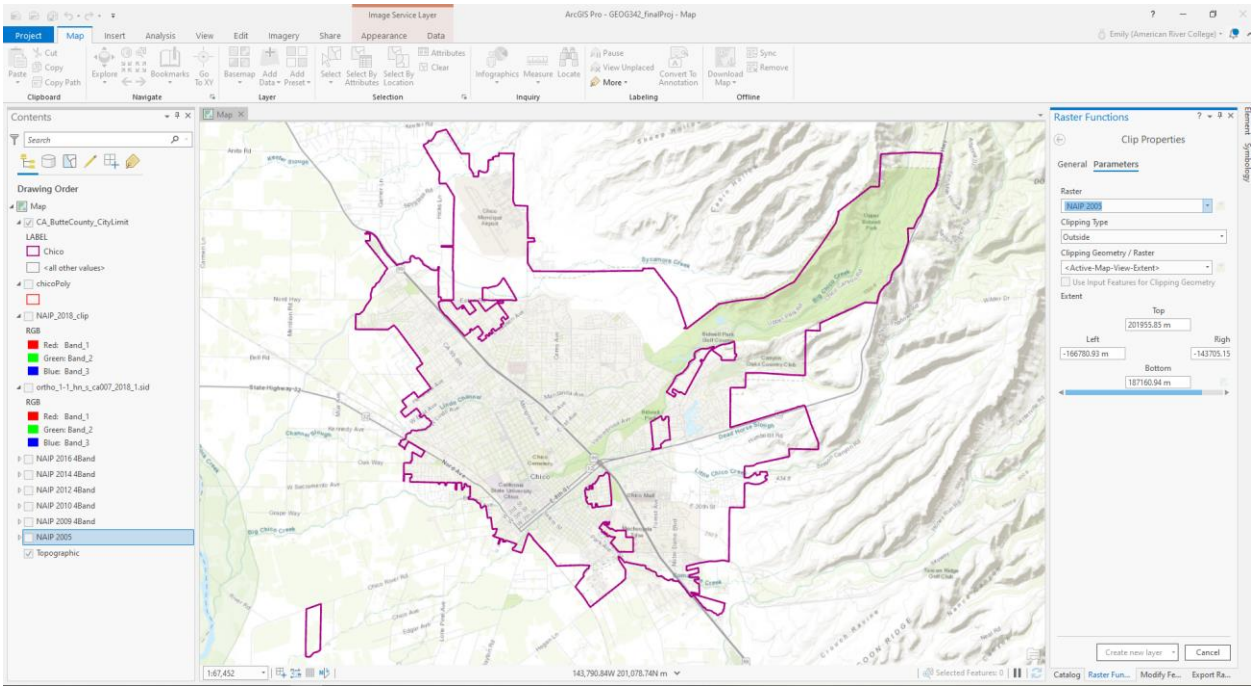


Figure 3. The chosen parameters for the output of this clip operation.

After clipping the desired layers and organizing the table of contents, the image segmentation begins – or so I thought, but I'll get more into this in the discussion section. For now, this is just my process with the hopes of what I might accomplish. For image segmentation, there is tool that is found by looking under the imagery tab. The desired image needs to be highlighted in the table of contents, and then under the "Classification Tools" is the option for "Segmentation."

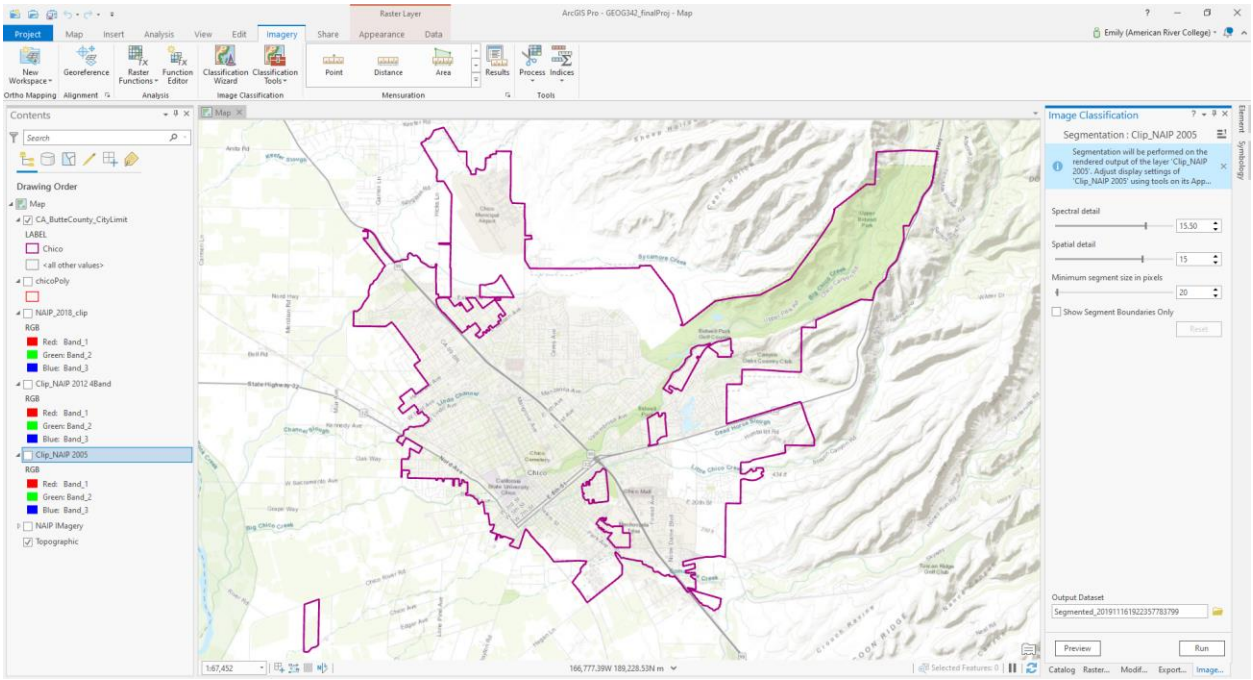


Figure 4. The "Segmentation" tool with the default settings when you open it in ArcPro.

I learned through trial and error that the defaults were a little but more than I needed (see discussion below), so after reading up on the "Classification Wizard" tool some, I decided to try my luck with some different settings. Before running the segmentation process, I saved my map with a different name in case something happened and I had to start over (and I cannot emphasize how important this became as I got more into this project!). After messing with the settings a little and viewing different areas of my map to see how it would look, I came up with an output that I felt was suited to my needs, gave it a more meaningful name, and ran the tool.

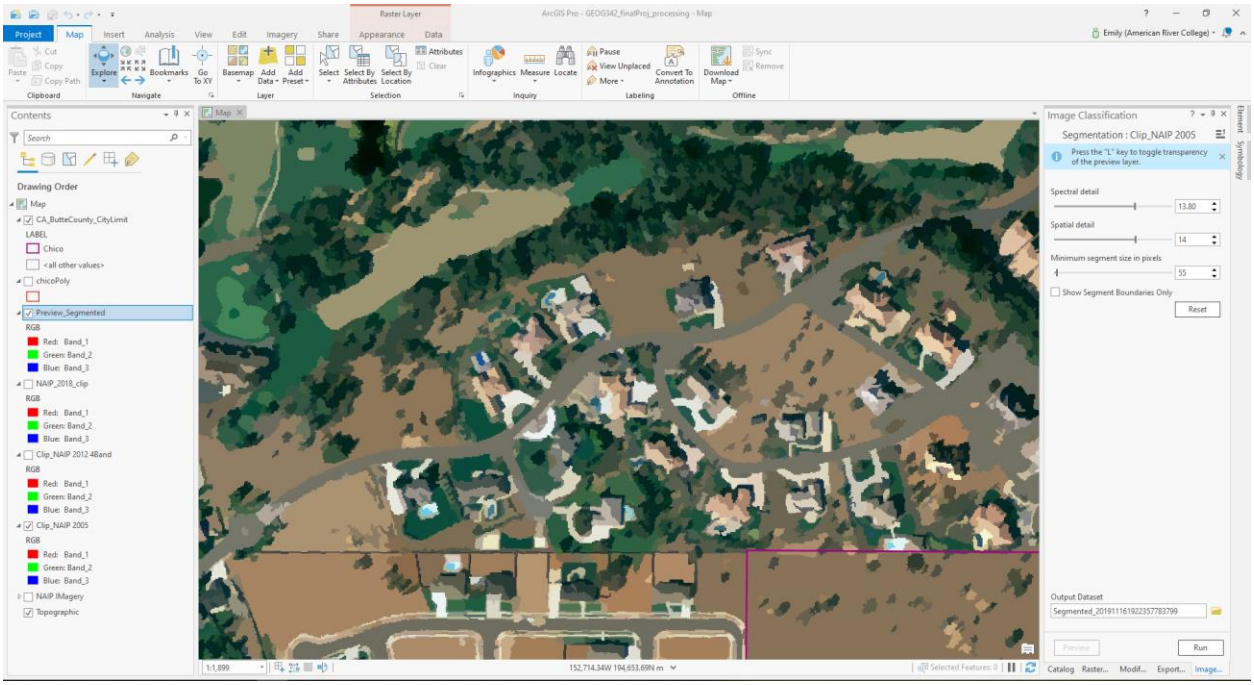


Figure 5. Use of the preview option to see how the segmented image will look once the process is done.

Once the tool finished running, I was to view the image in full, and spot check some trouble areas I had in my first trial (not pictured). The surrounding areas of Chico have a lot of orchards, as well as trees in Bidwell Park, and oaks up in the foothills, and I was hoping to be able to get my image to differentiate between them.

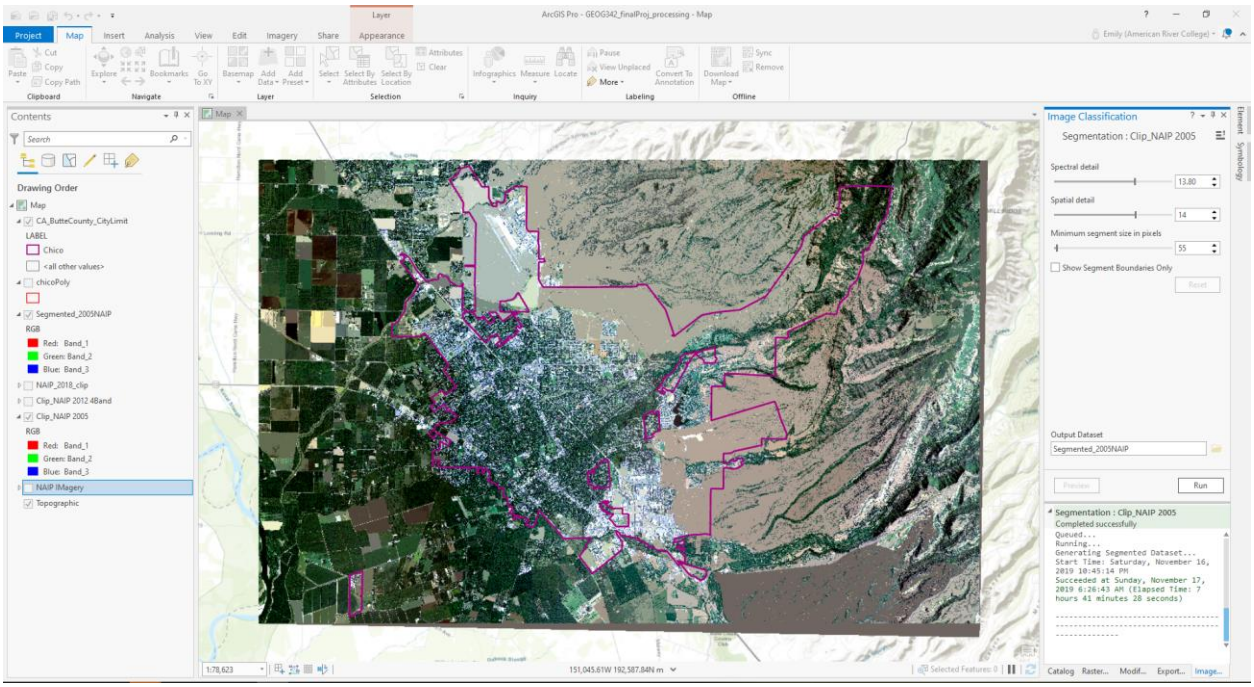


Figure 6. The finished result of the "Segmentation" tool with the desired settings.

At this point I was satisfied with my result in terms of features I was looking for were broken out in the segmented image (it only took roughly 8 hours to get), so I selected the layer in the Table of Contents and opened the “Classification Wizard.” I had already made a class template in my trial run (not pictured) so I loaded that into the wizard, and started

clicking some samples. I was a little disappointed with the result of my segmentation since the trees were still coming up in small patches, and some of my swimming pools were also eliminated. I zoomed in and counted up the pixels for a few of the pools, and one was at 27, the other at 40, so I decided to rerun the segmentation tool with a smaller pixel setting to see what I could do. However, before running it, I made sure to zoom into one of my "problem pools" to see how it would look after the segmentation, and I used the "Show Segment Boundaries Only" option in the preview to make sure that my pool of interest would be included.

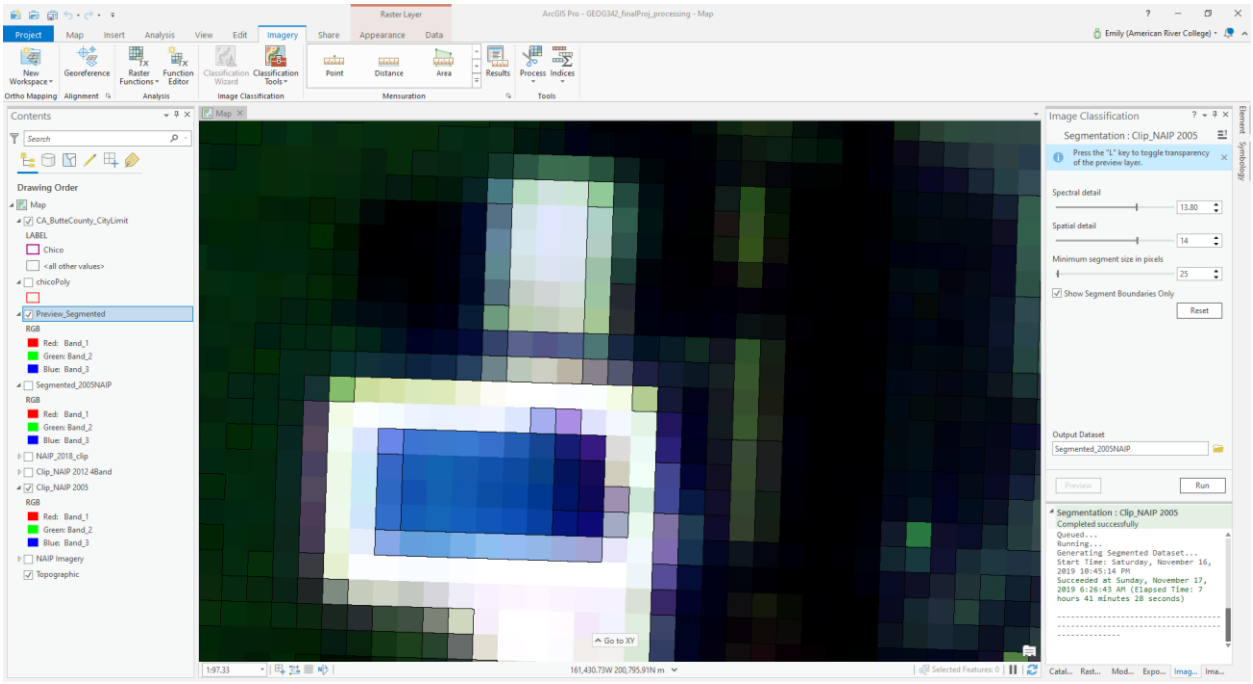


Figure 7. Zoom of one of the "problem pools" that disappeared in my classification.

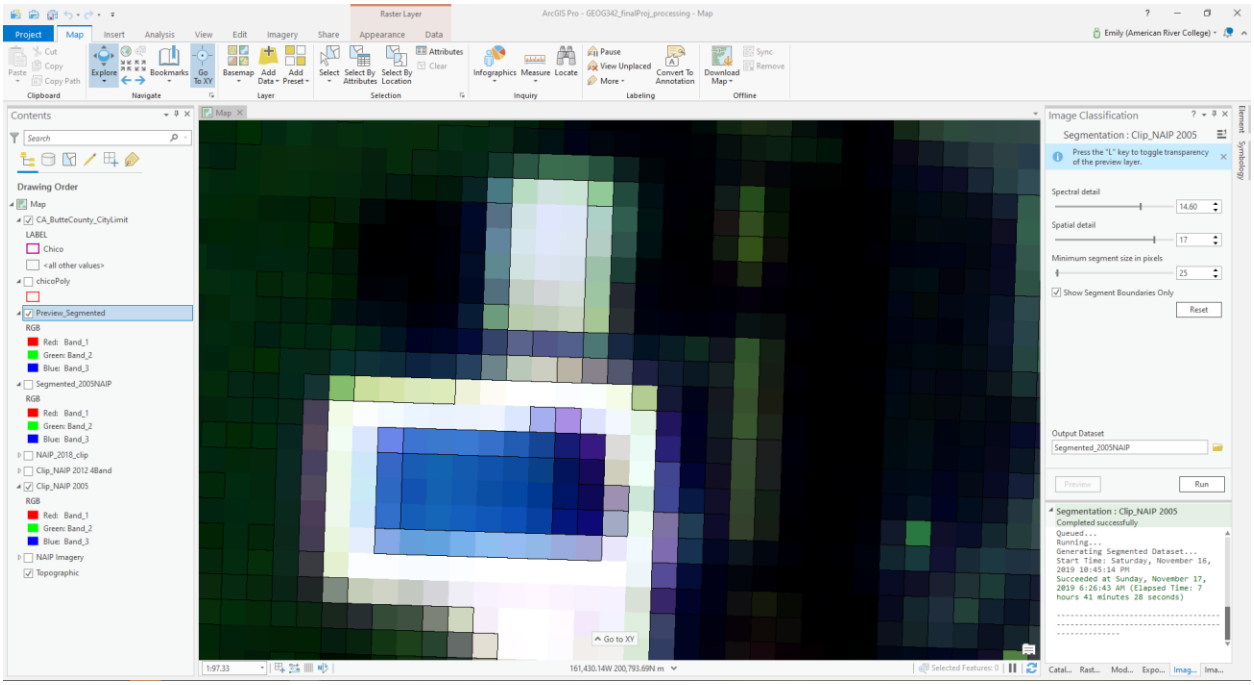


Figure 8. Problem pool with a slightly higher level of spectral detail, a higher level of spatial detail, and a minimum segment size of 25 pixels to ensure that smaller pools in the image were included.

I also wanted to make sure that my roads weren't becoming too indistinguishable from my barren areas, since I had noticed in my trial run that in some cases where the pavement was more worn, it would merge into the "lawns" (many in this image were barren earth, or mostly dead). Once I played around with the settings, I gave it another try. I was heartbroken when my computer crashed, and 12 hours later I had nothing to show for it. It was at this point I decided it was best to take a pause in the project, and talk to the pros to see where I was going wrong. I had my suspicions, but I was so frustrated that I figured it might just be time to ask for help.

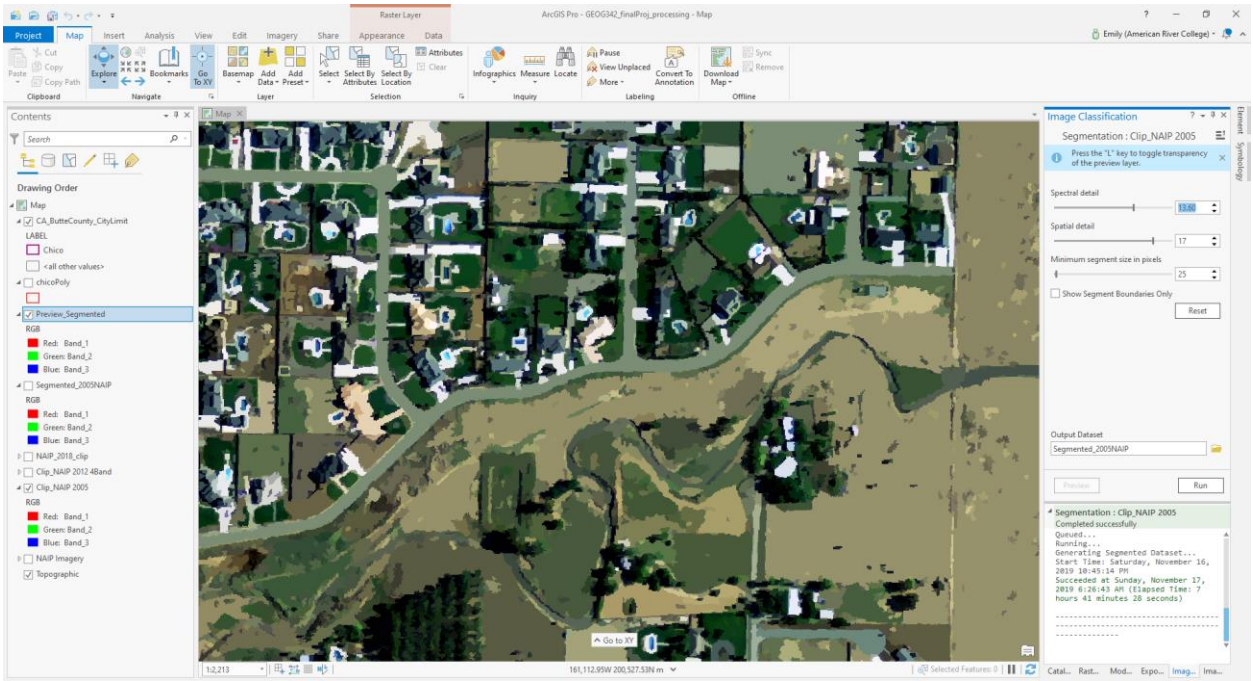


Figure 9. At this level of spectral detail in the segmentation tool, the roads and some barren areas can blend.

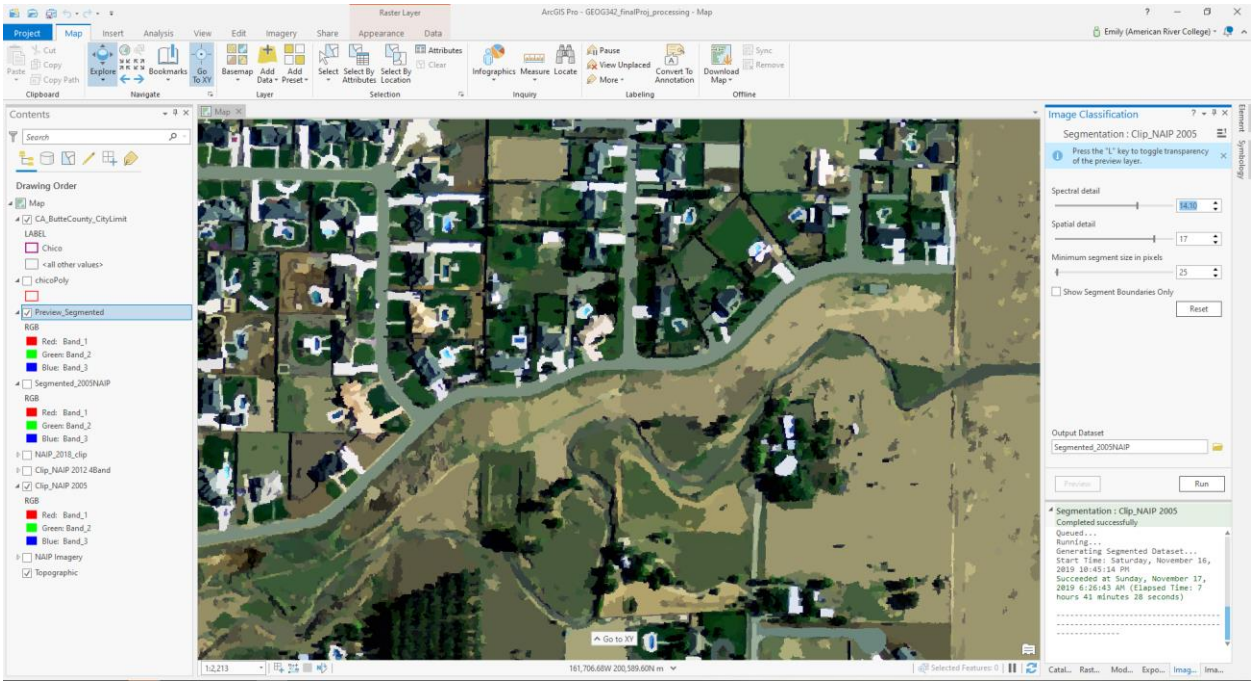


Figure 10. This mix of spectral and spatial detail allows the roads to be more distinguishable from the barren areas.

Unfortunately, this was as far as I was able to get with my 2005 imagery, for a multitude of reasons, including the lack of an infrared band, but I'll save these issues for the discussion section of my paper.

2018 Imagery

After learning all that I thought I could learn about this process the hard way with the 2005 imagery, I thought I might be ready to tackle the 2018 imagery and at least see if I could still get the classification to pick out a few of my features of interest. The following images document the trials and tribulations of this process, and I will say it was most definitely a learning experience.

From my crash and burn experience with the 2005 imagery, I learned that I needed to export the NAIP after clipping it to get the tools to work with it. I also opted to try this on a smaller area of Chico, rather than the whole city, just to see if I could get it to work. Since I was still interested in seeing if the classification tools would pick up the swimming pools, different types of trees, and various developed features, I picked an area that I knew was more likely to have a mix of those features. So, I started with the 2018 NAIP that I had downloaded for work a few months ago from the NRCS website. Once I had my area of interest, I clipped the image using the methods described above, and then I exported it to a .tif, which I neglected to do on the first round. I attempted to do this with the other datasets I had so I could attempt to get a comparison still, but since I was using a webhosted layer, I was restricted in the size I was able to export, and didn't have the time to track down other NAIP imagery for the area. Given all the issues I was already having, I figured from this point forward, it was best to keep things simple so that I could at least see if I could get the software to work for me.

Through my many experimental runs, I learned that if I wanted to pick up features like swimming pools, I would need to use higher spectral AND spatial details when I began the segmentation process. Using a spectral detail of 16.5 and a spatial detail of 15, with a minimum pixel size of 25 (I kept the number of pixels in swimming pools that I had counted earlier in mind for this), I was able to produce a segmented image that I was happy with, and it didn't take the better half of a day to complete.

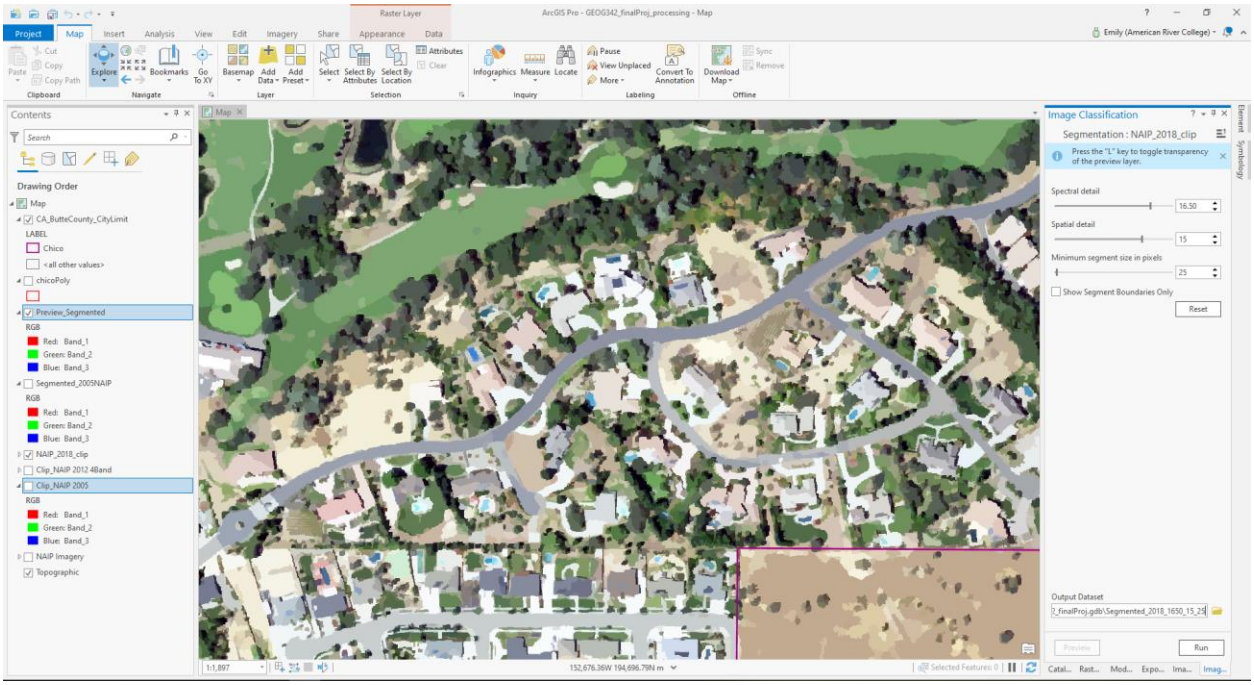


Figure 11. 2018 was a much drier year than 2005, so it was necessary to use higher values for the spectral detail so that the roads were distinguishable from the barren areas, and so that pools were not lost to the concrete.

Classification Wizard

Now that I had usable clipped and segmented images, I was ready for the "Classification Wizard." This time I selected my clipped 2018 NAIP image, then added my segmented image and picked my custom classification I had created earlier, and was ready to go. I tried all sorts of combinations, and got a different result each time. In some cases, I wouldn't pick enough barren areas and so random things would be classified as barren. Chico is mostly oak and grey pine in the foothills, but landscaping has a variety, and portions of Lower Bidwell Park were in my image and there's a mix of species in these areas, which required me to select a LOT of samples with my custom classification method. Rooftops were incredibly varied, so I attempted to get a large sample population of these as well. And the pools… I wanted to make sure I had enough pools but even if you click 100, they're such a small portion of the training classes I wondered if they would ever work out. Since this was NAIP imagery and I had selected such high quality for the output segmented image, EVERYTHING was in tiny pieces, even rooftops and tree canopies, so clicking all those parts took some time. The following set of images document the samples I took with my custom class, and are followed by the results I obtained in my previews.

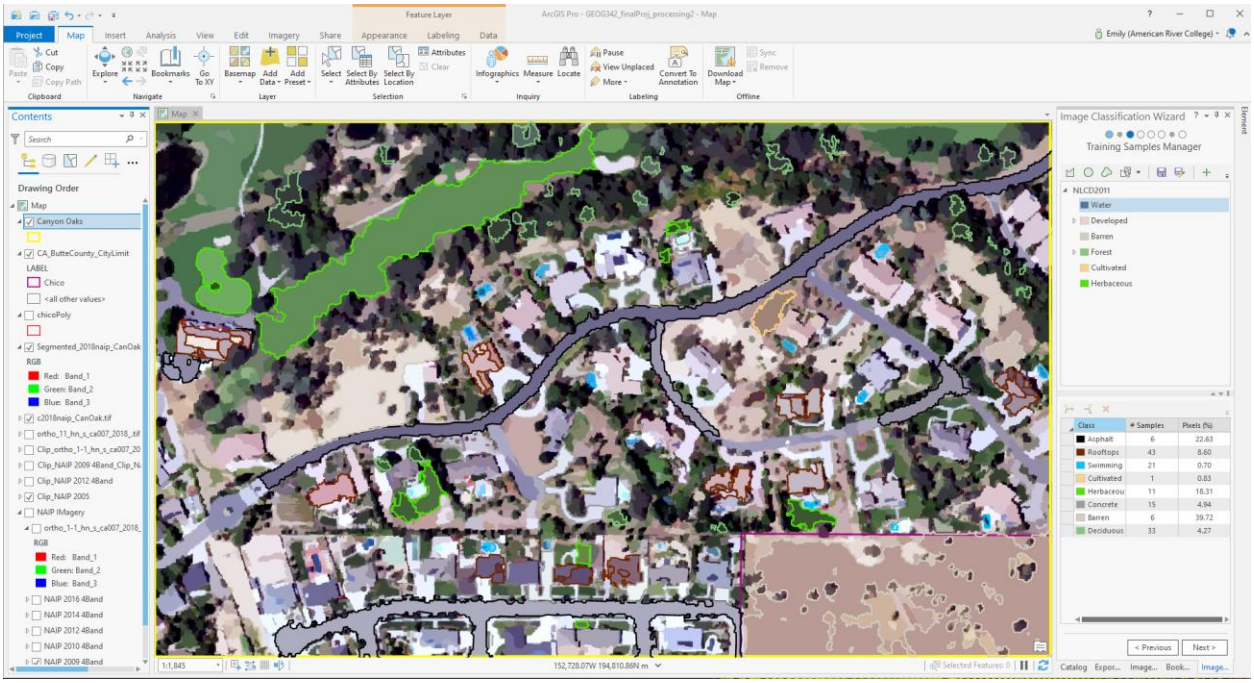


Figure 12. Sample area with training samples shown, and the modified classification system I had developed earlier in the process. The categories in the "Classification Wizard" reflect those selected in this sample. More were selected through the project area after.

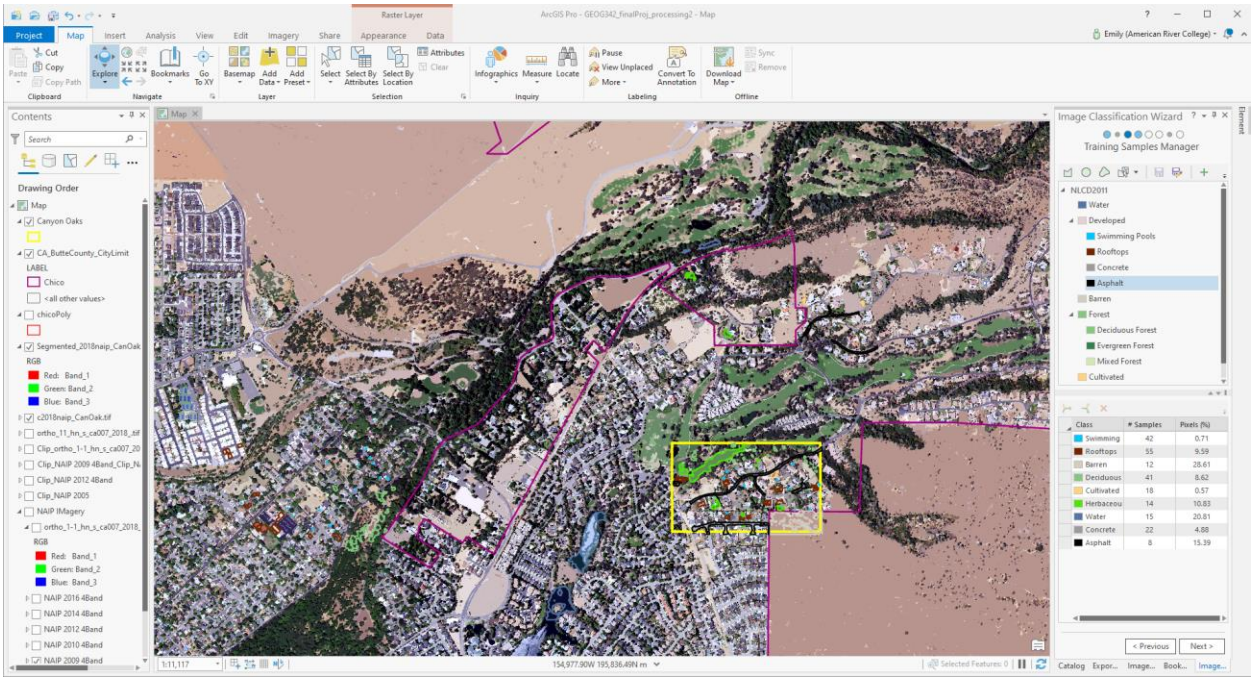


Figure 13. First try with the 2018 clipped and exported .tif image. The area with the yellow box around it is where the screen was zoomed to in Figure 12, and will be a point of reference from this point forward.

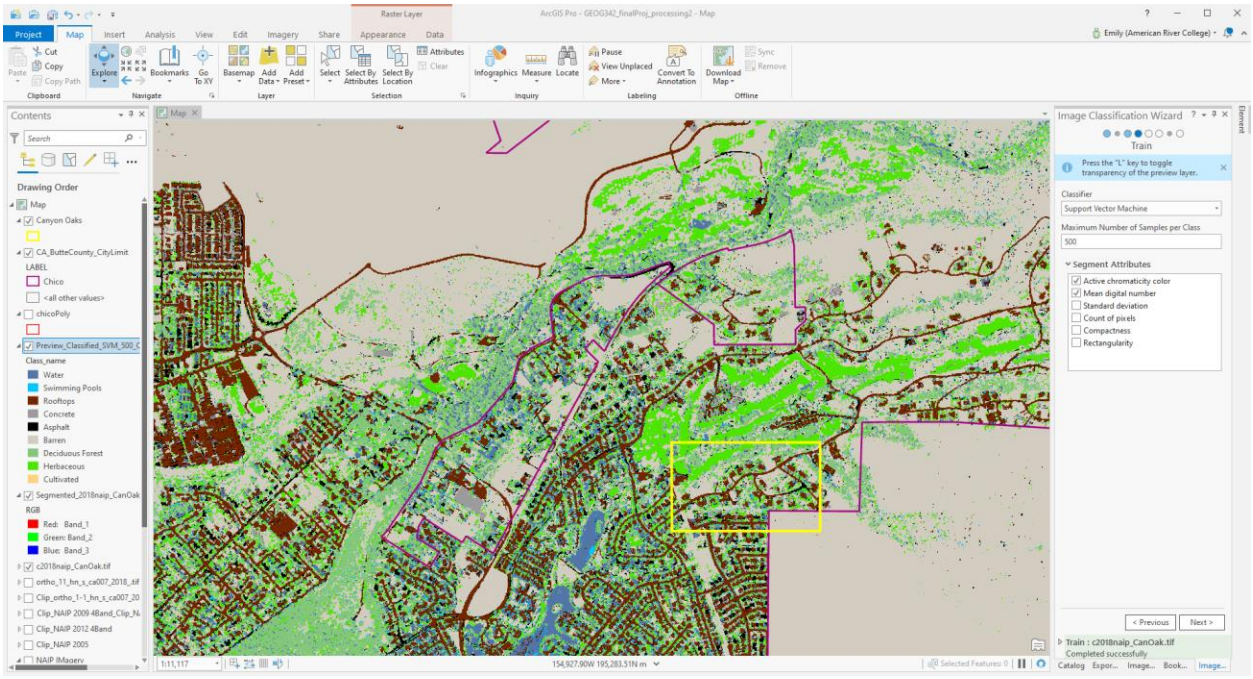


Figure 14. Definitely hitting the developed areas, but not quite picking out the types I was hoping to get. Roads are being classified as "rooftops," and rooftops are being classified as "pavement."

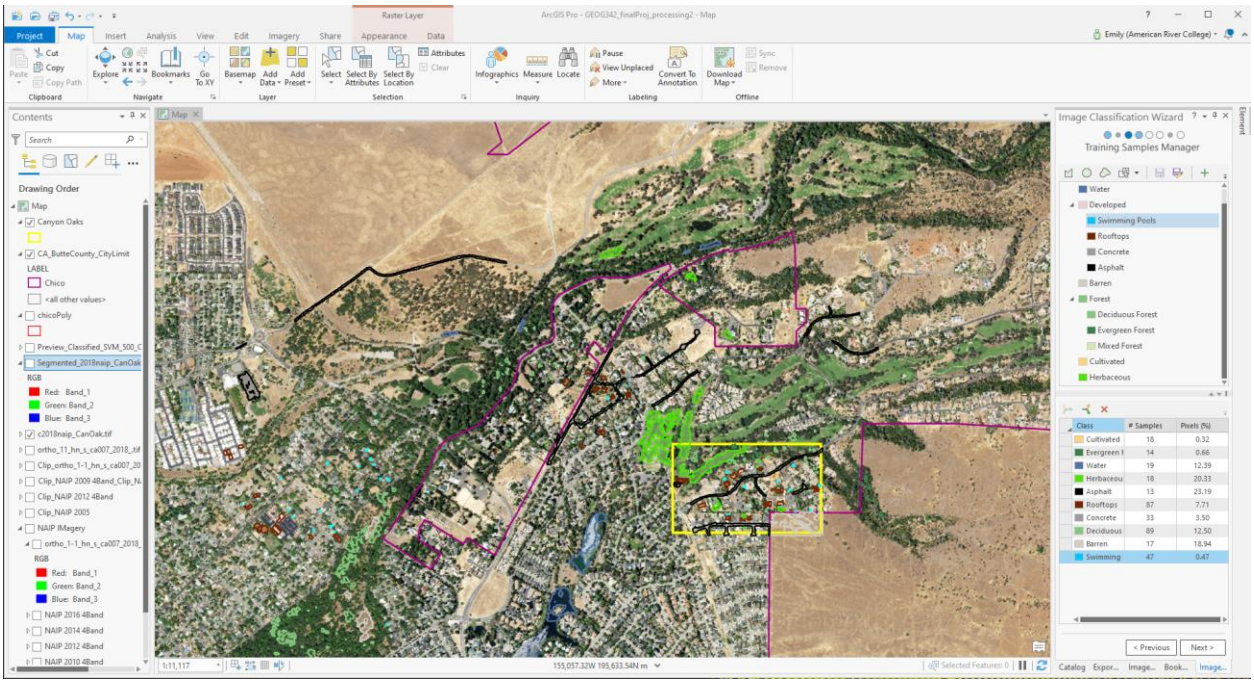


Figure 15. Second try with my custom classification method. Though the pools seemed to pick up pretty well, I added a few more samples just so that they didn't get lost.

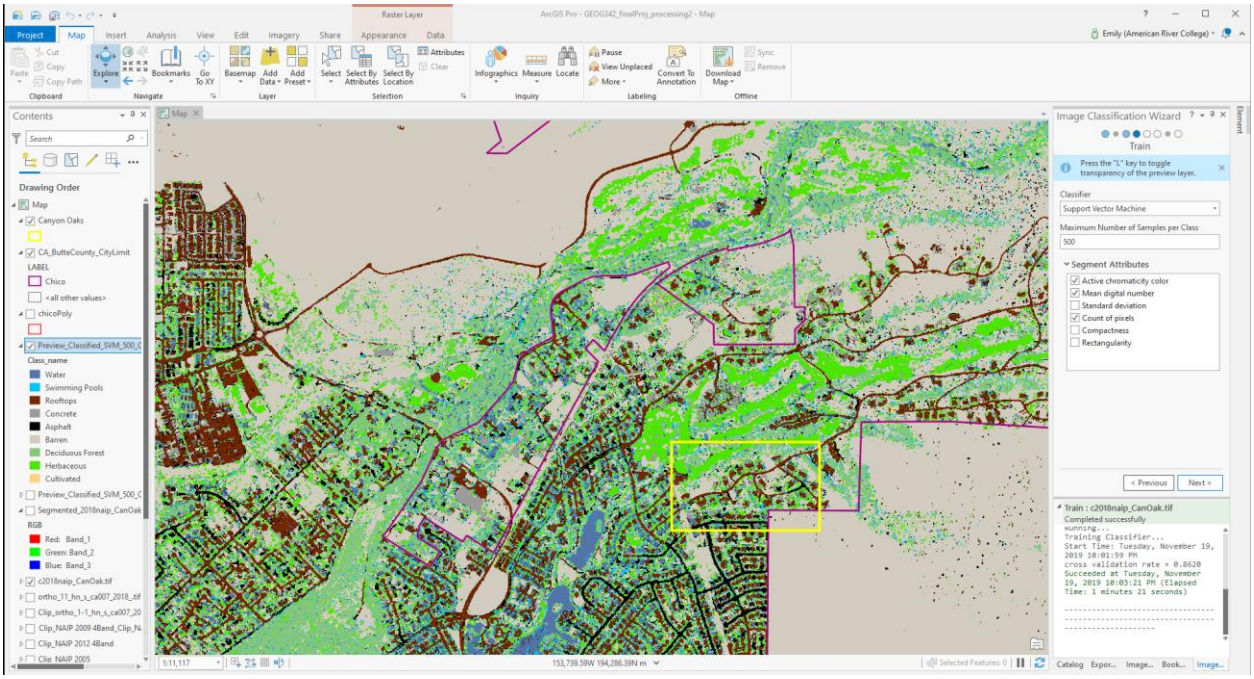


Figure 16. Result of second training. Some trees are still showing as water, and the creek in the northern portion isn't picking up despite there being a few samples in it. Interestingly, some roads that have been selected as "asphalt" are showing up as "rooftops."

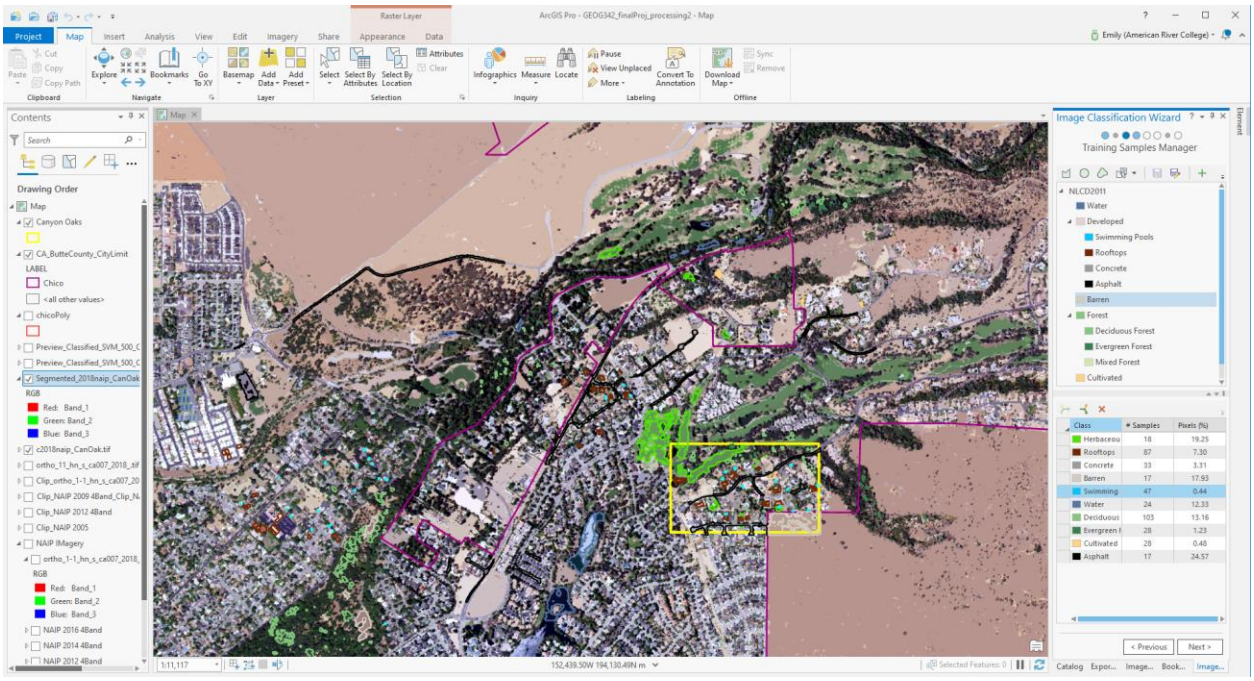


Figure 17. Third attempt at differentiation between developed features using my custom classification system (shown with segmented image).

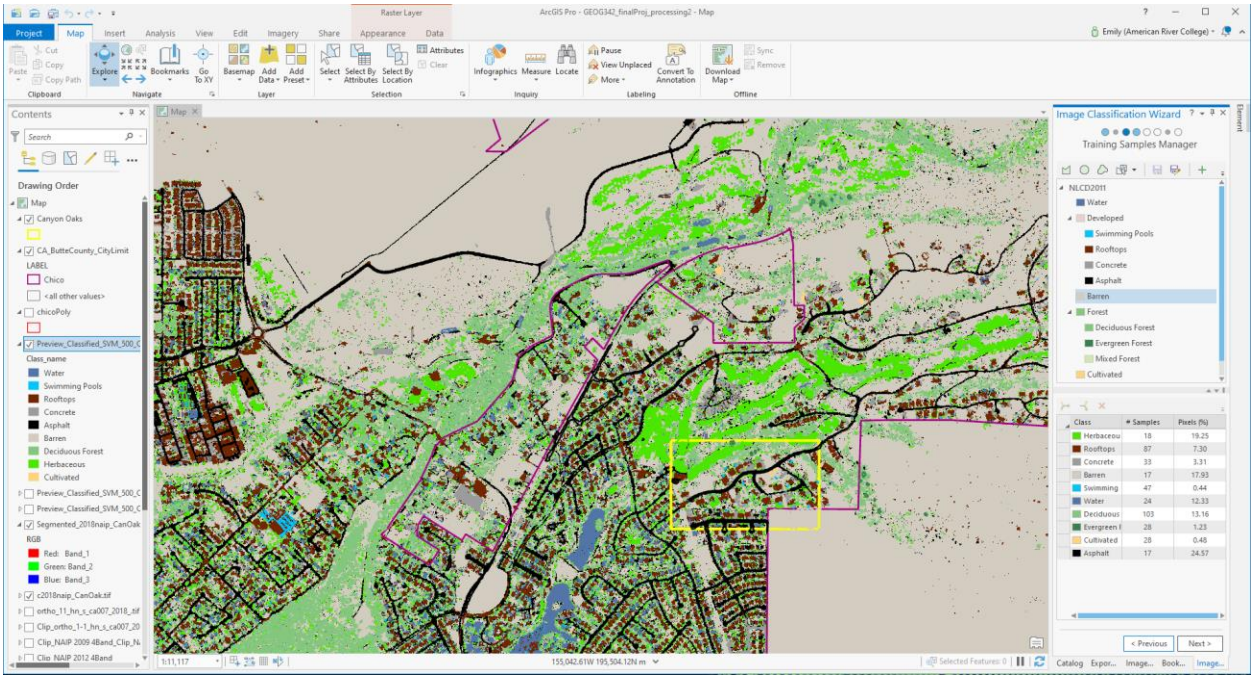


Figure 18. This still left some tree shadows showing up as water, so more trees were added to the training. The streets, roofs, and pools are showing up nicely though!

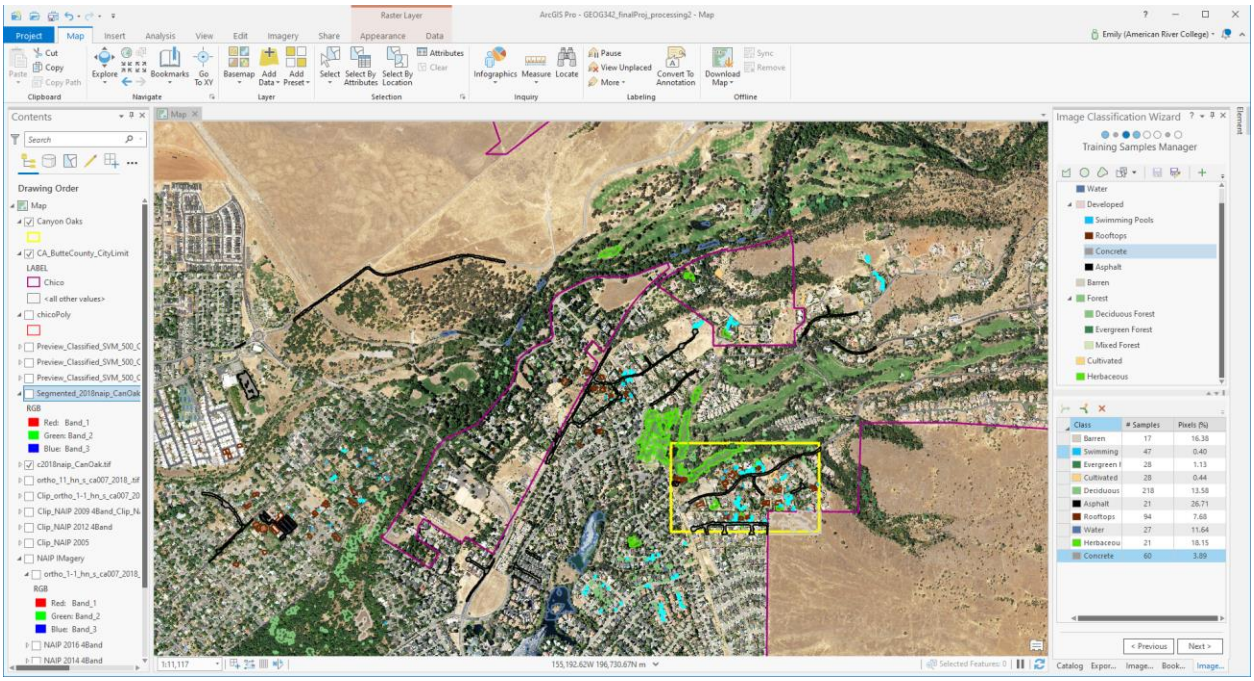


Figure 19. More trees (and their shadows)! More water, more concrete, and go!

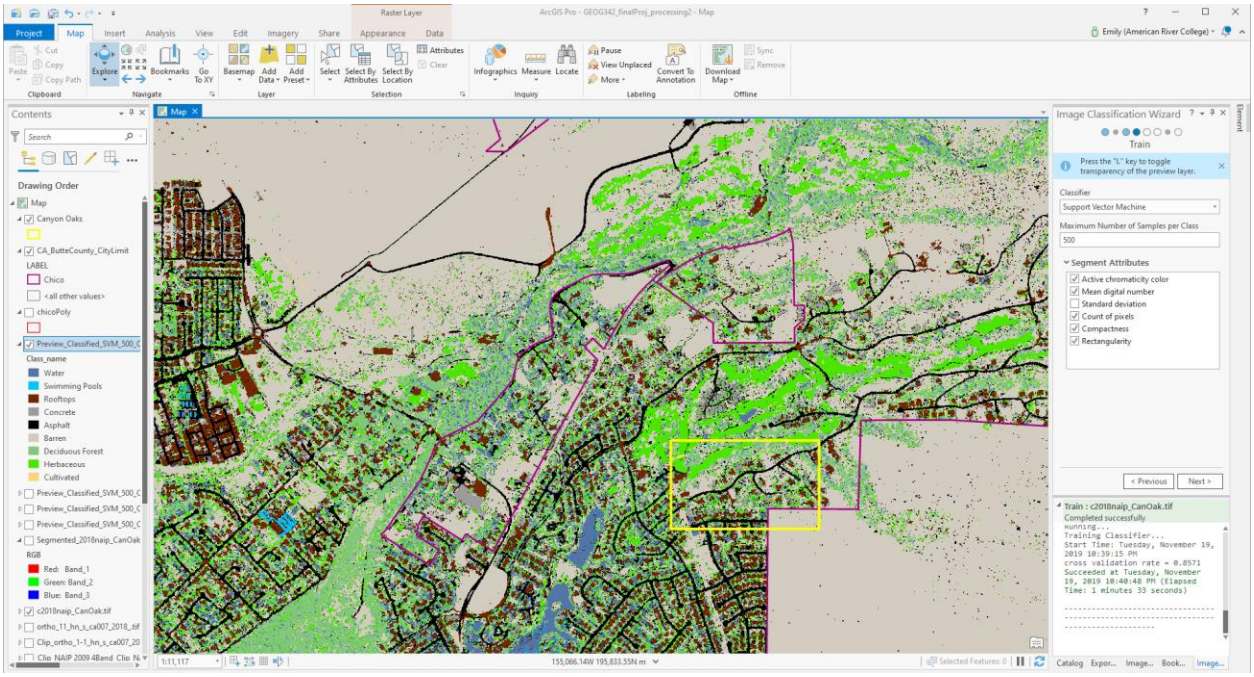


Figure 20. And with more samples, more turns into water, not just the tree shadows...

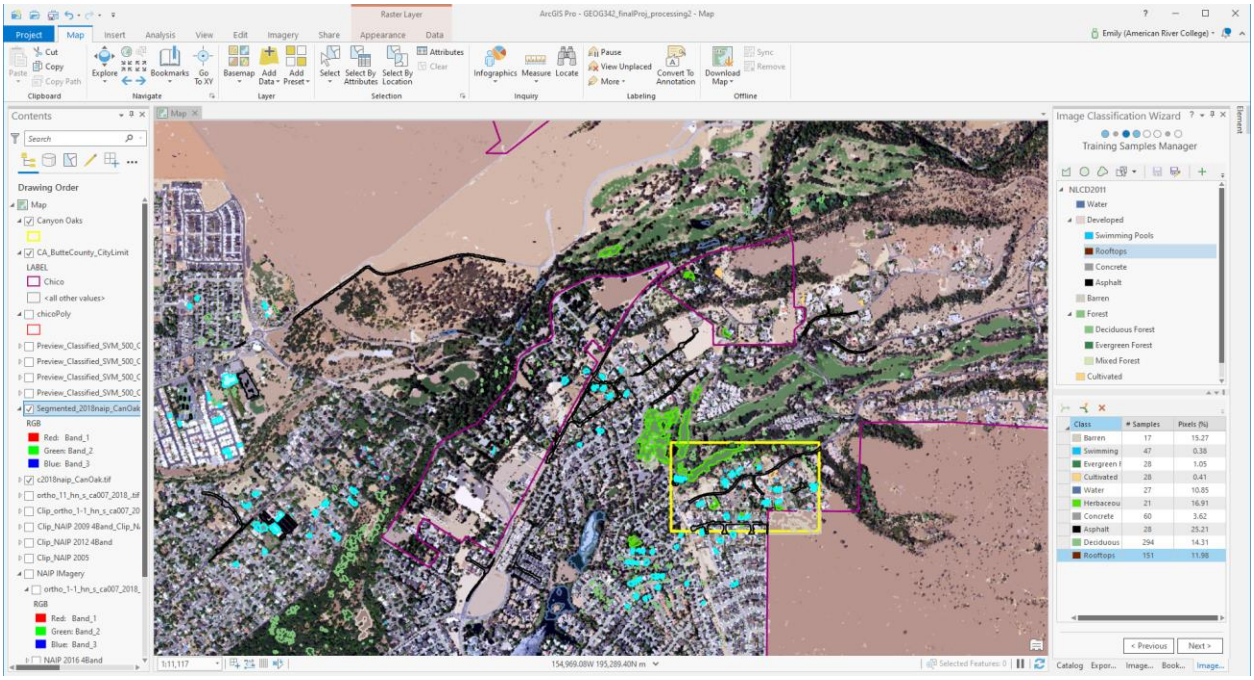


Figure 21. So, I tried adding more samples...

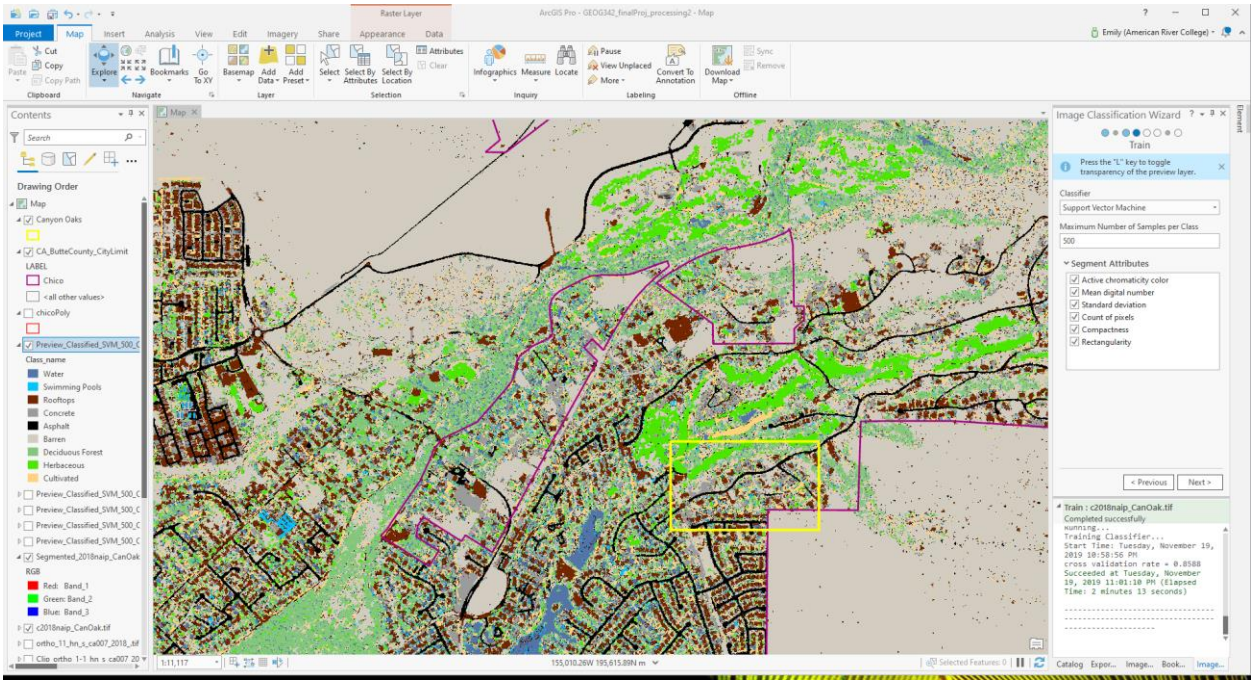


Figure 22. And then I got this weird influx of cultivation. There were really only two spots that I saw in this area that were cultivated, and I had only selected them both to see if something else picked up.

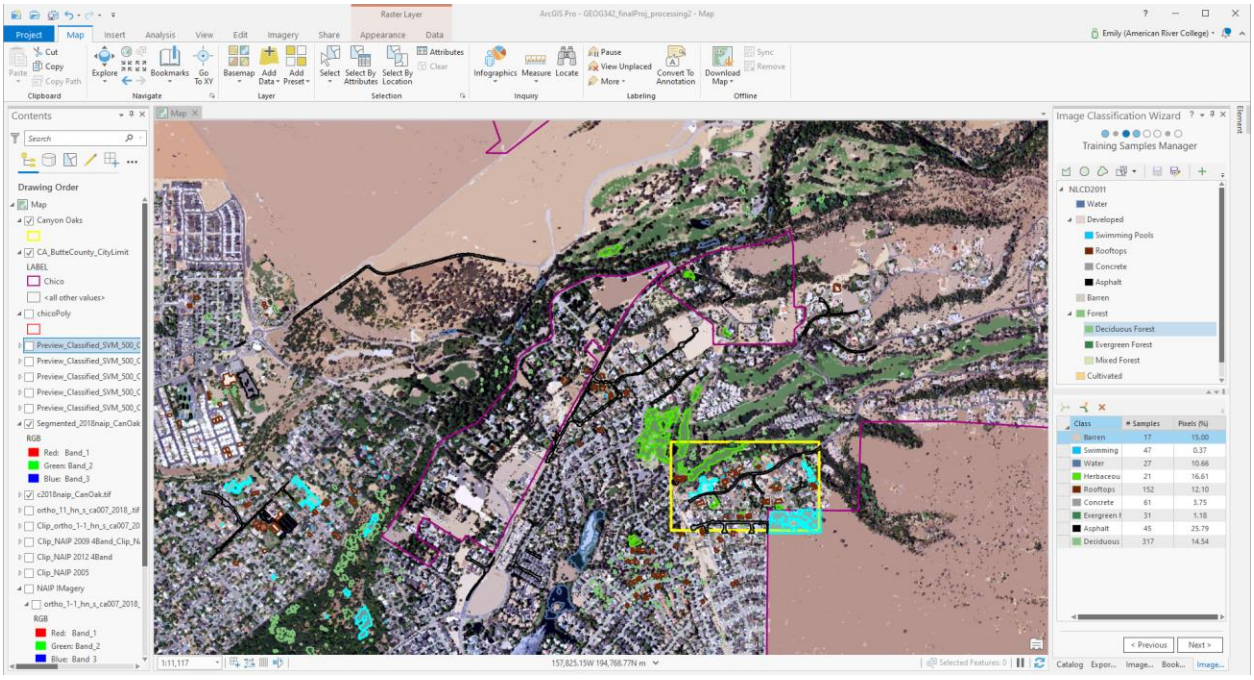


Figure 23. I ended up removing cultivation since it was such a small portion of my image, those areas were probably orchards anyway. I added more trees (with shade), roofs, and concrete.

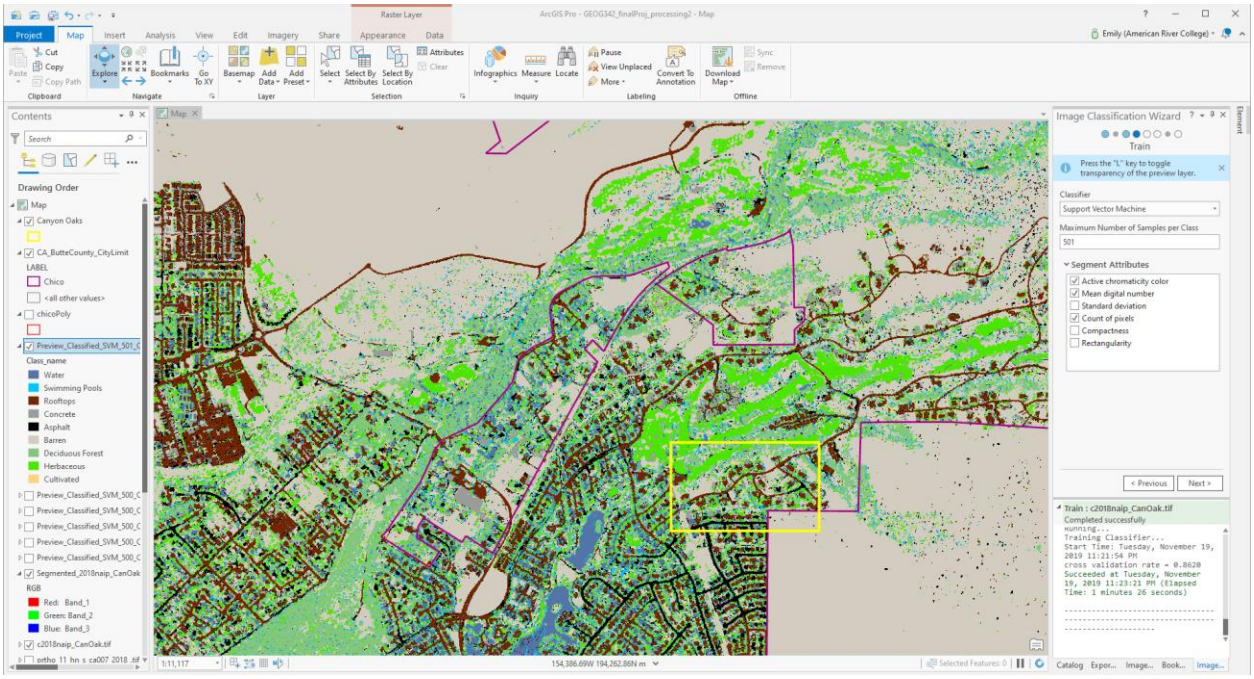


Figure 24. With more samples, I began to introduce more error. I had more roads coming in as rooftops now, and rooftops coming in as water.

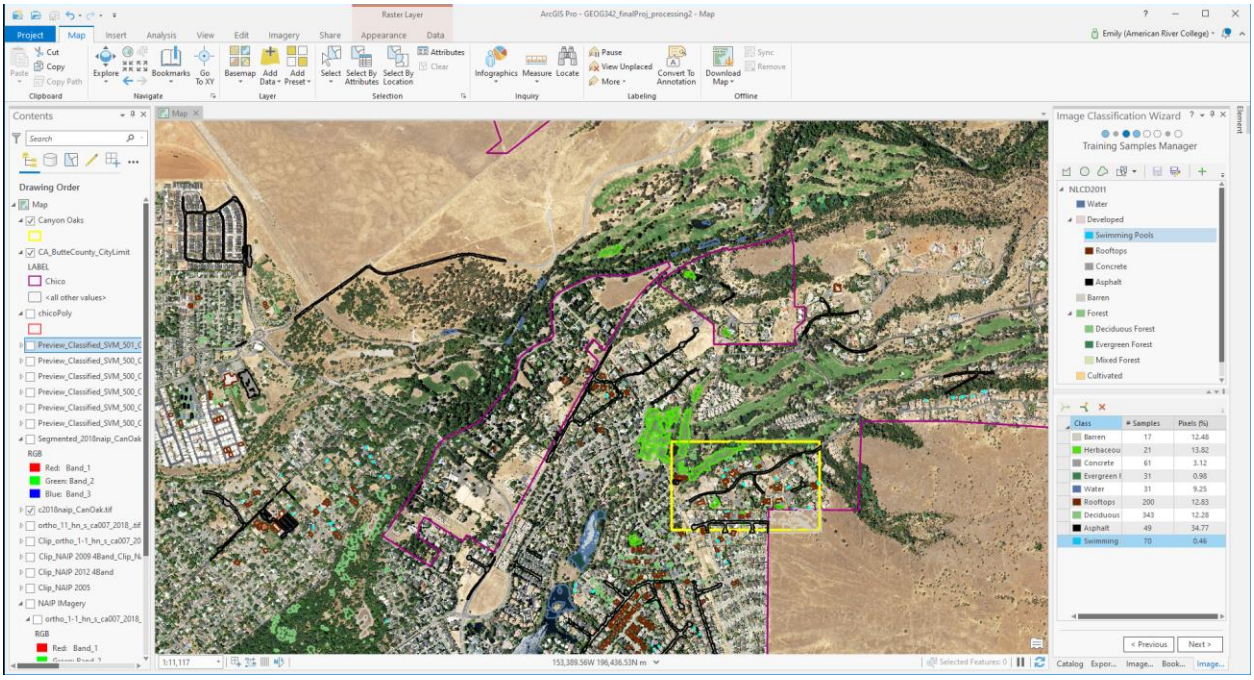


Figure 25. I still tried more samples, thinking maybe I was just deficient in the areas I was having issues with, rather than considering that it was possible I had too many.

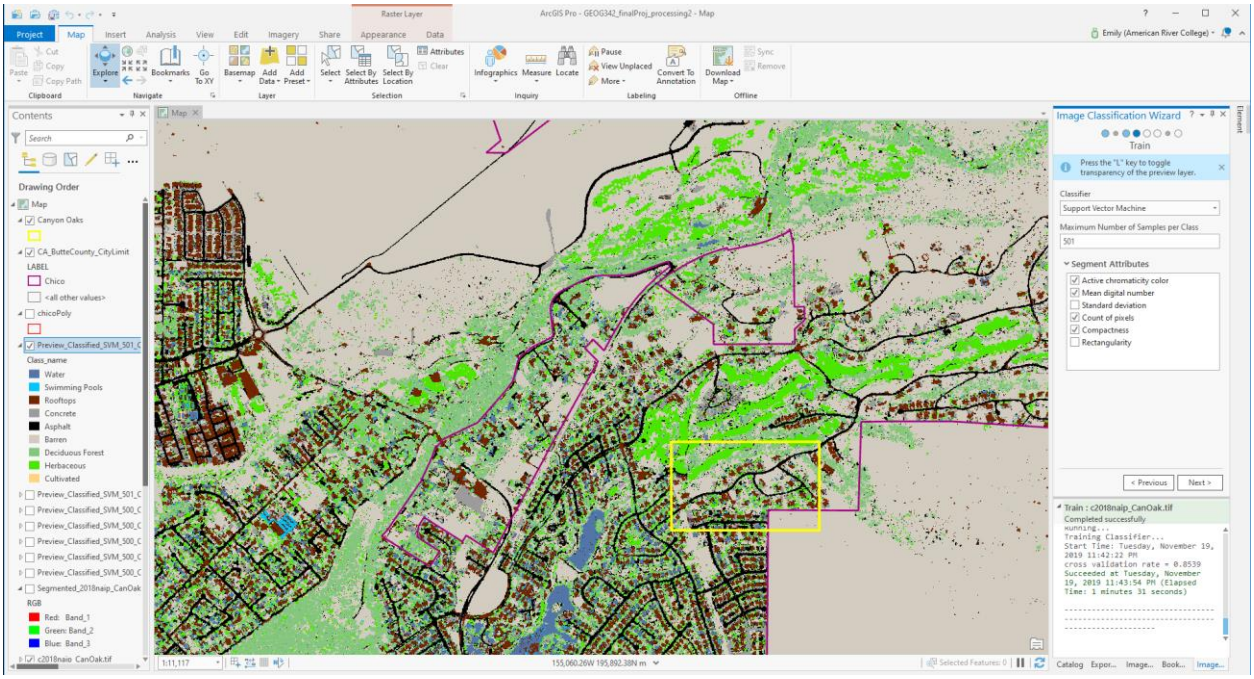


Figure 26. After the last round, better, but… still quite a few roofs showing up as water, and the creek not showing any water in areas not trained.

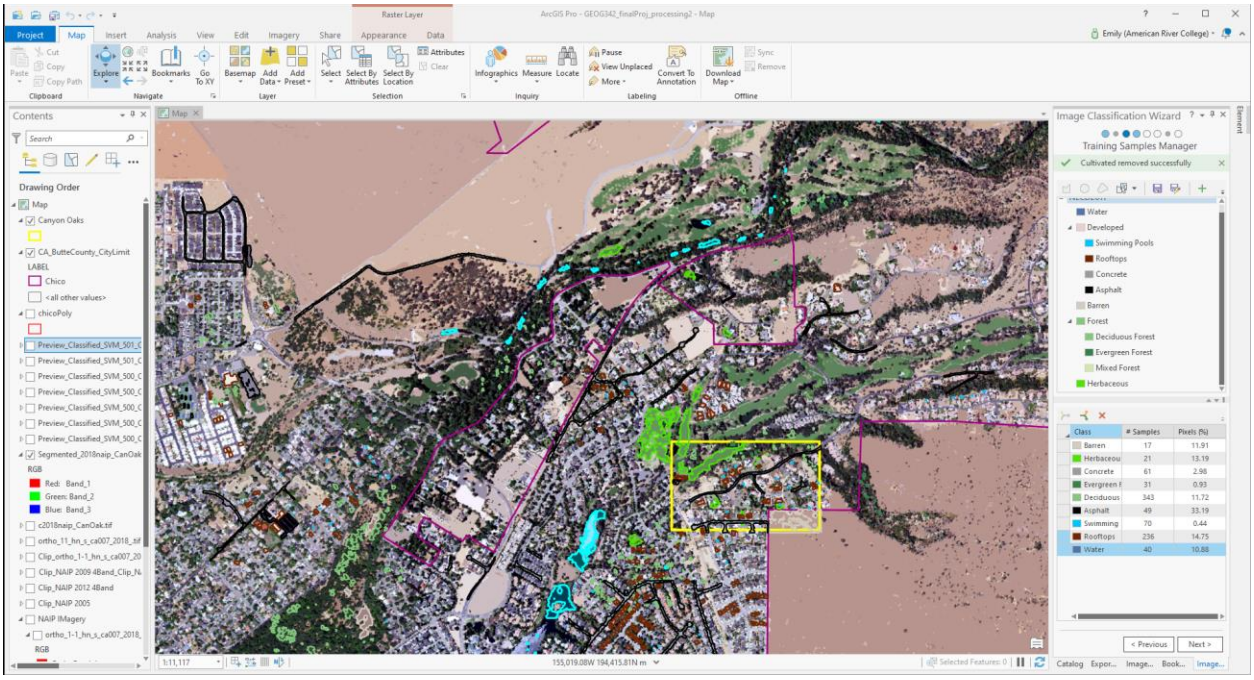


Figure 27. So, I tried even more samples…

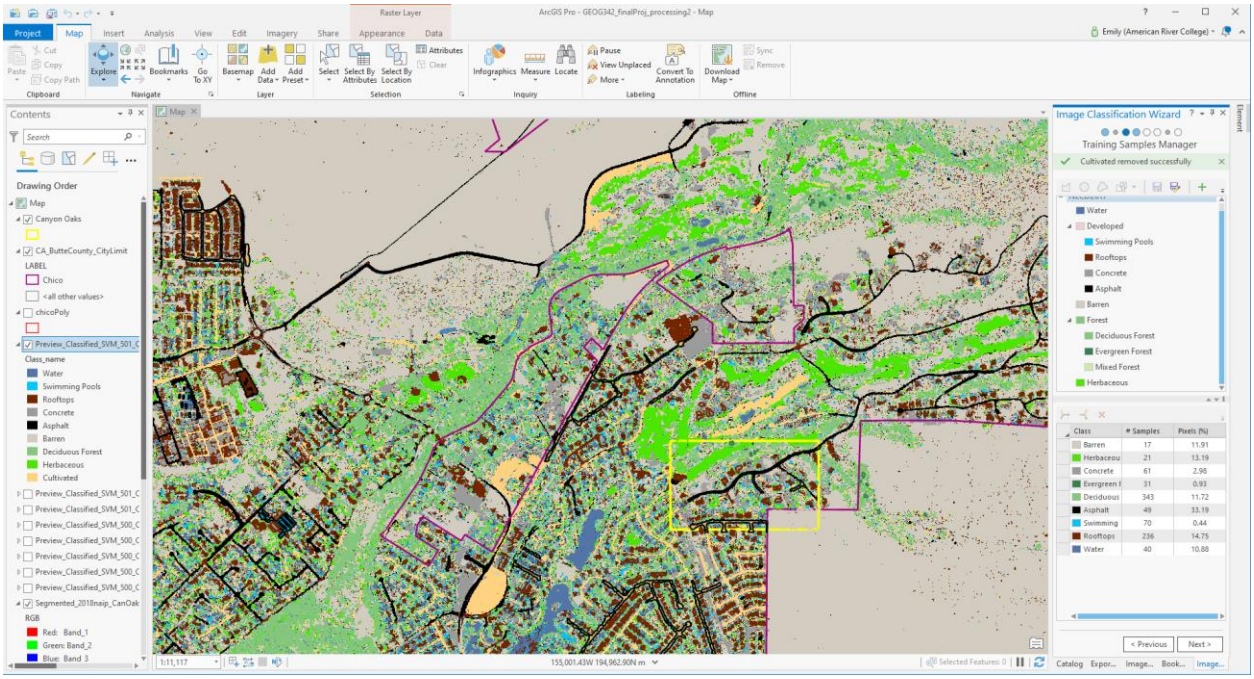


Figure 28. Cultivated, despite being absent from the map, has made a triumphant return. I'm guessing because I shouldn't have removed it at the time I did. Rather, I should have maybe left my original samples, and dealt with it during a reclassification.

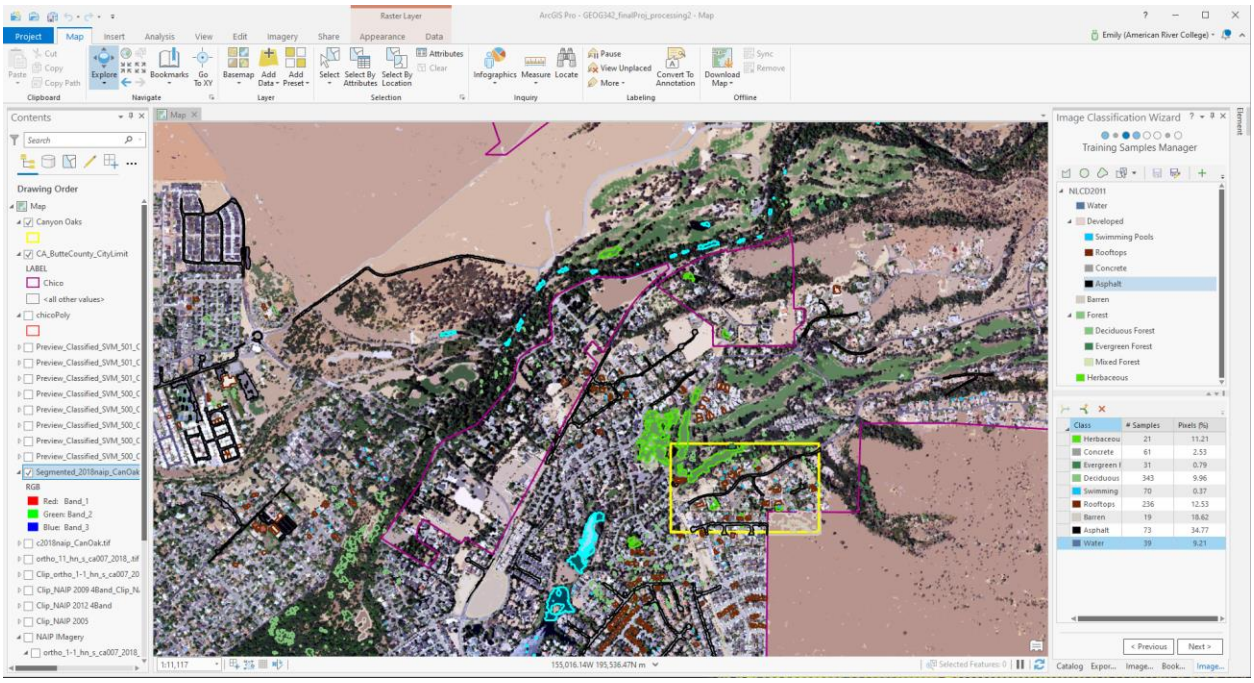
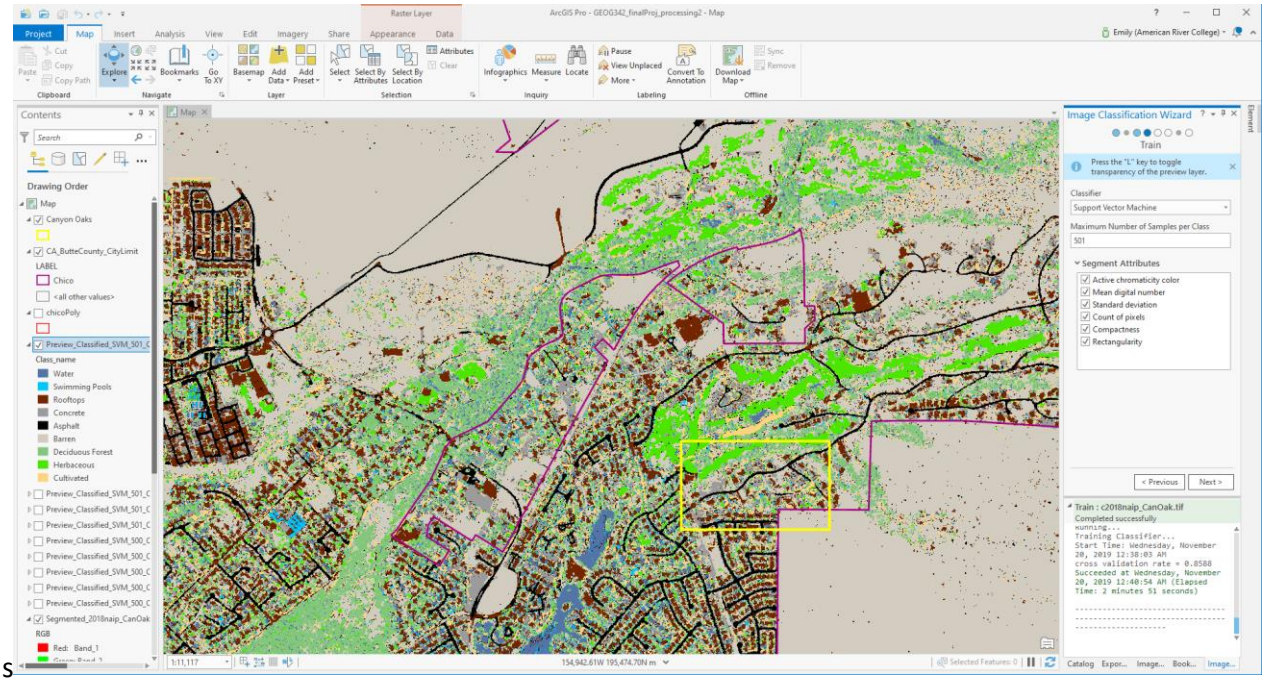


Figure 29. So, I tried even more samples. I still thought that maybe more was better.



S

Figure 30. This got me closer to what I wanted. Still not perfect, but I was ready to push forward!

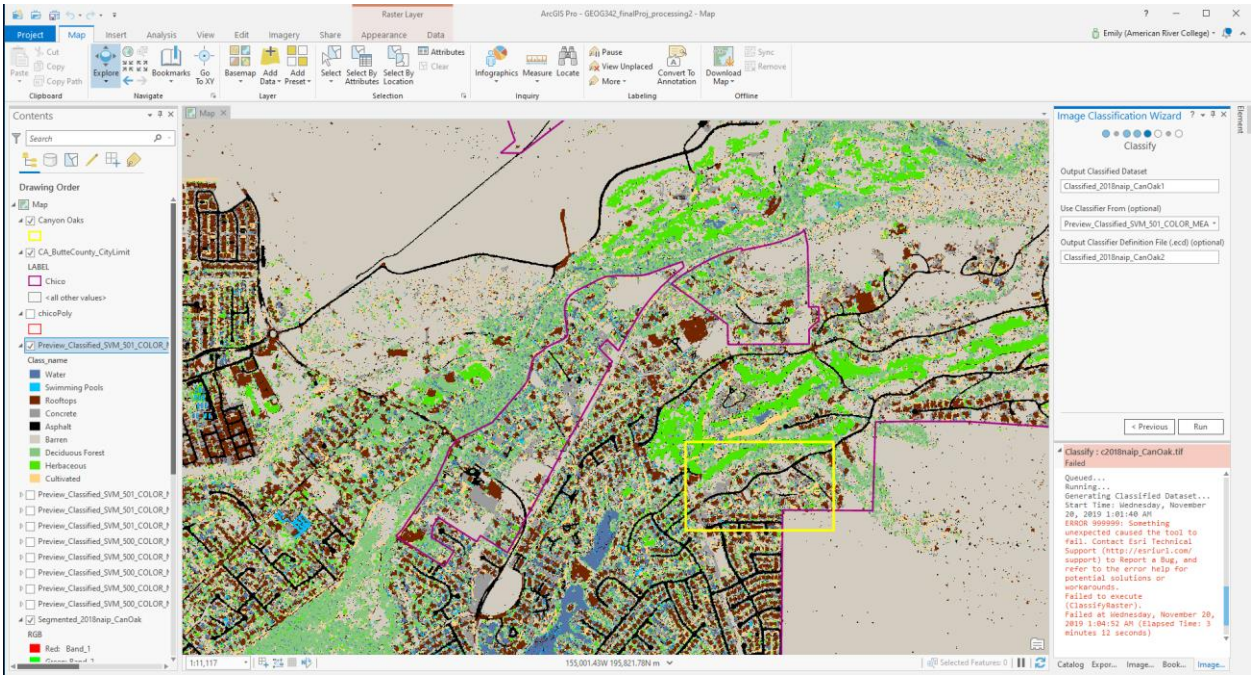


Figure 31. After 2 trials, the classifier would not run with this preview. My best guesses are that this could have been related to my messing with my classification scheme, the number of times I went back and added new samples, the number of categories I was trying to classify, or just that ArcPro can be a bit of a jerk sometimes.

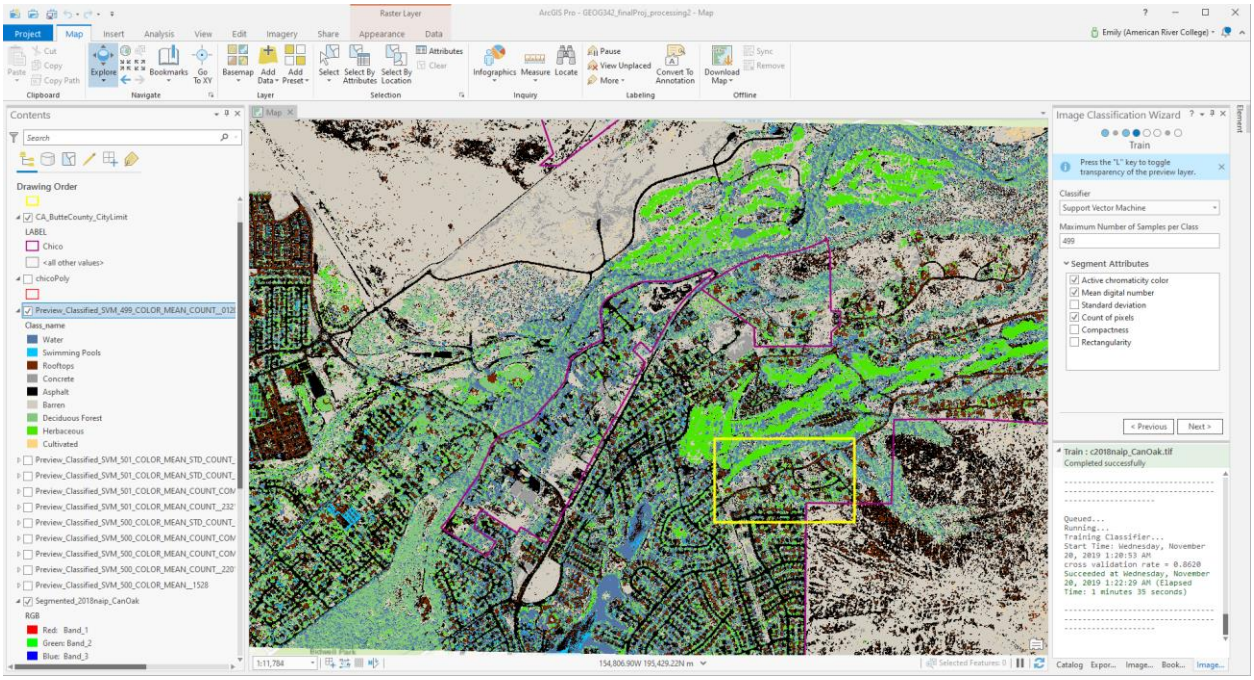


Figure 32. After several failed attempts of moving forward trying to use the other previewed layers I had created, I finally broke down and Cultivated fields were added in, and the classifier was run again.

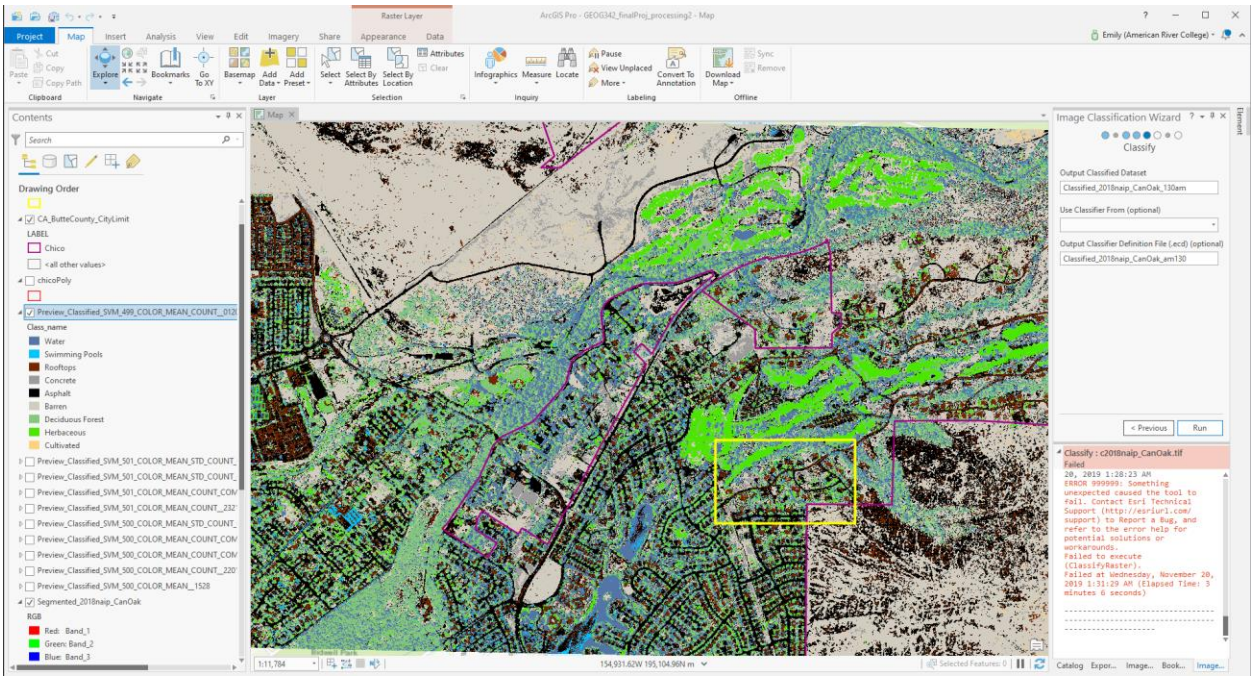


Figure 33. And again, the classifier would not run.

After losing as many hours as I did with the above, I realized it was probably time to try and take it down a notch, and use less classes. I wasn't ready to give up my search for swimming pools, but I did concede to removing the different classes I had been trying to pick out in the developed areas. I also decided to just look for "forests," and dropped the differentiation between coniferous and deciduous trees. I totally removed the cultivated field so I didn't have to deal with it, and was ready to give it another try.

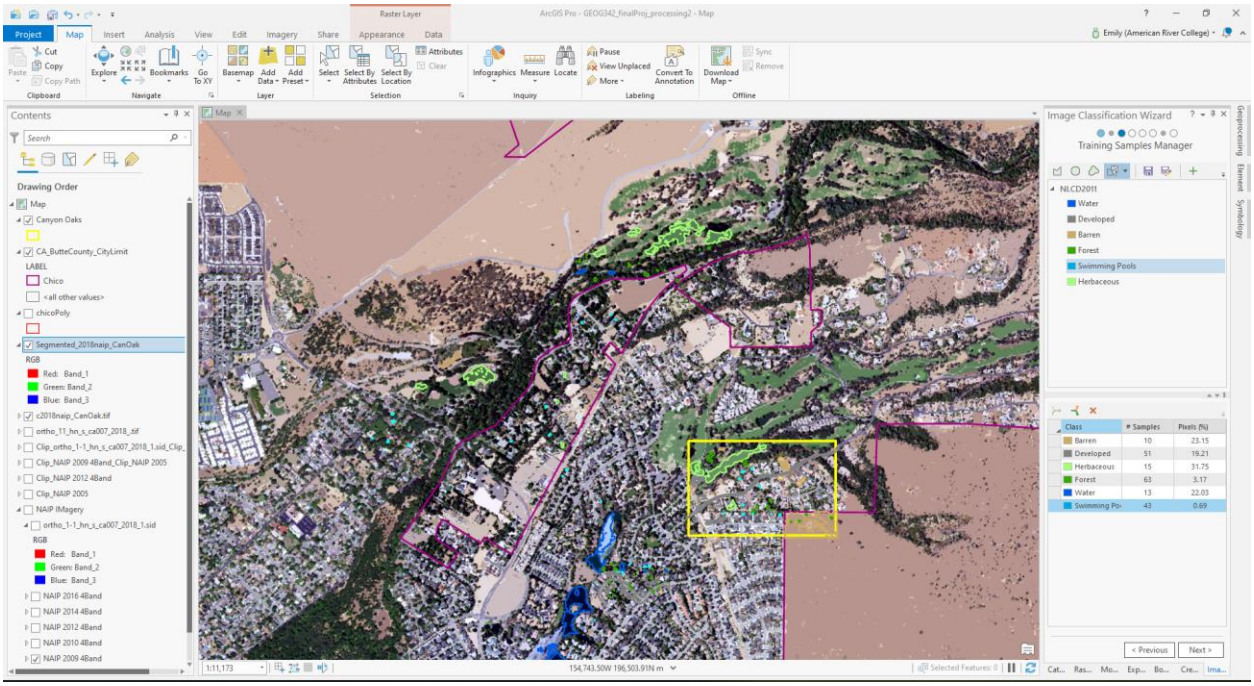


Figure 34. On this attempt, I used the same segmented image I had created earlier for the 2018 NAIP, and cut back on all the excess… except for swimming pools. I wanted to retain some portion of my original project.

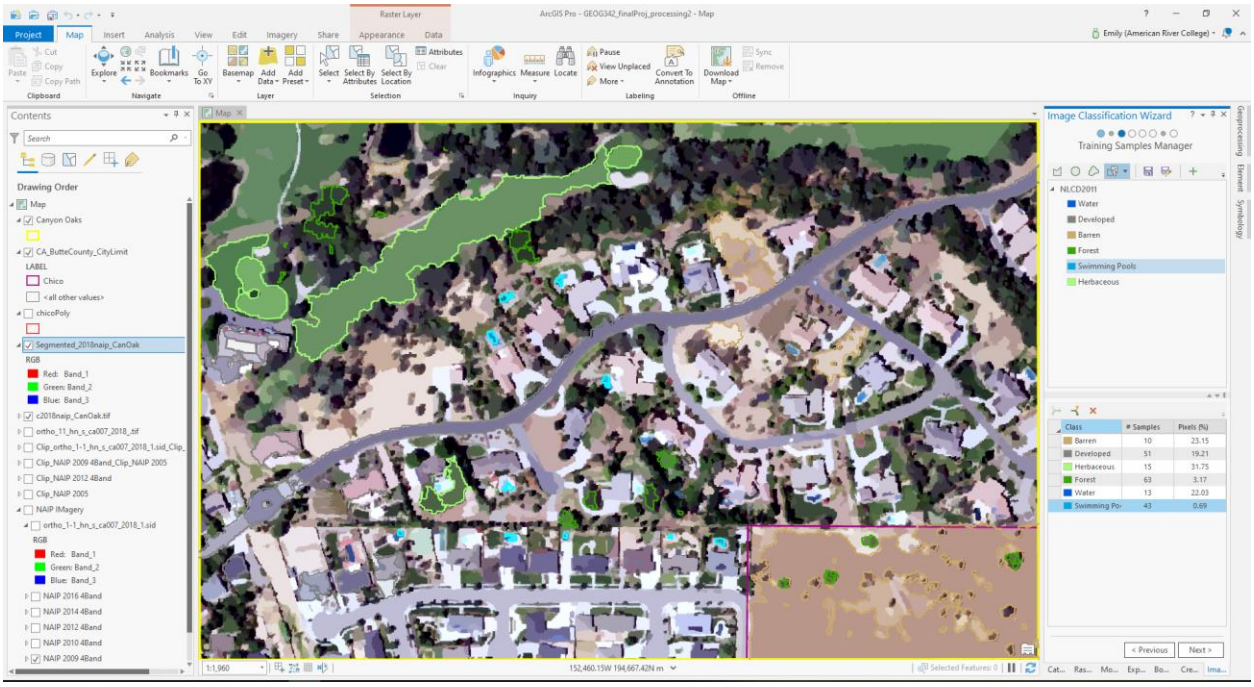


Figure 35. I selected some samples from the Canyon Oaks neighborhood that I mentioned earlier in this paper, and left part of the neighborhood unsampled so I could use it in a later comparison.

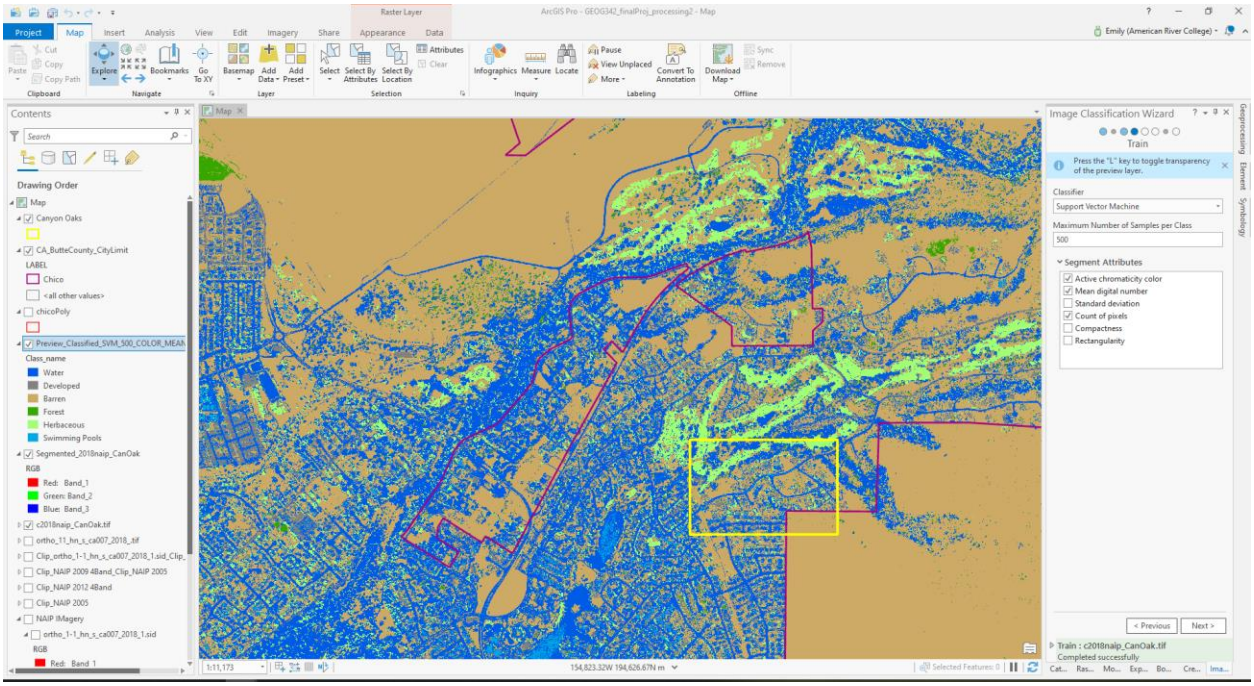


Figure 36. On my first try, everything was water.....

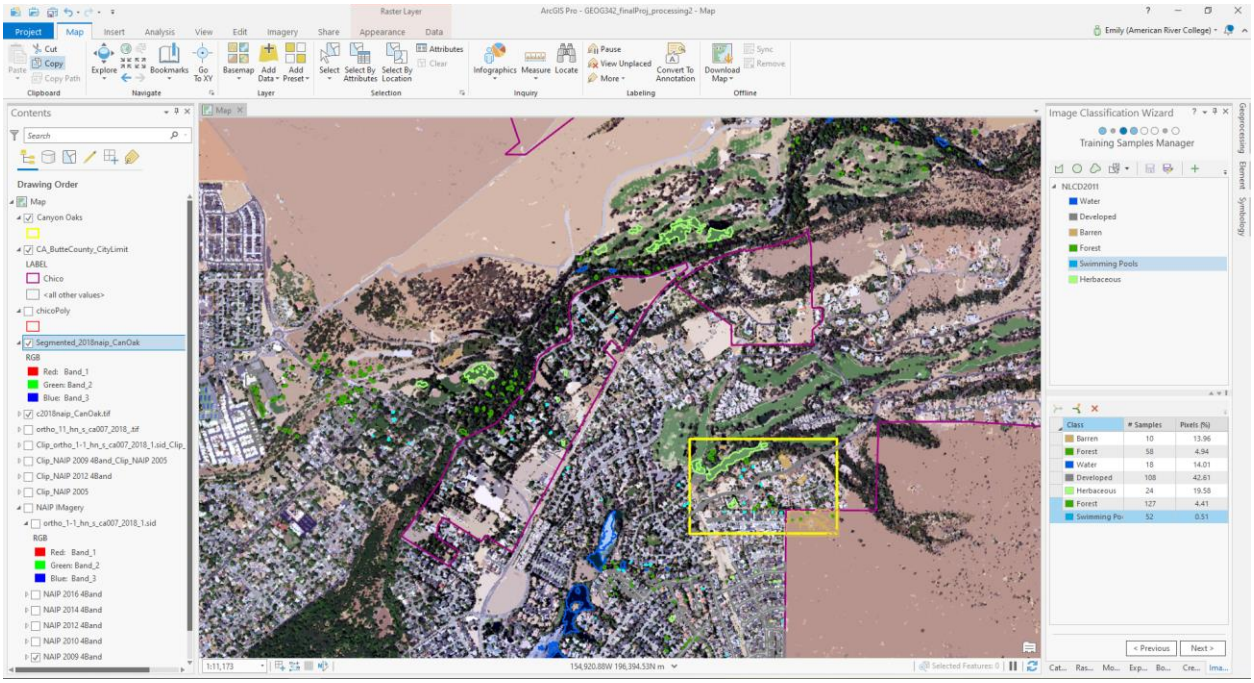


Figure 37. So, I went back, and I did a few more samples. I was hoping to avoid having to do several hundred like I had earlier in the process, but recognized that in some cases, it might be necessary, given the way things like trees and rooftops had broken down in the segmentation.

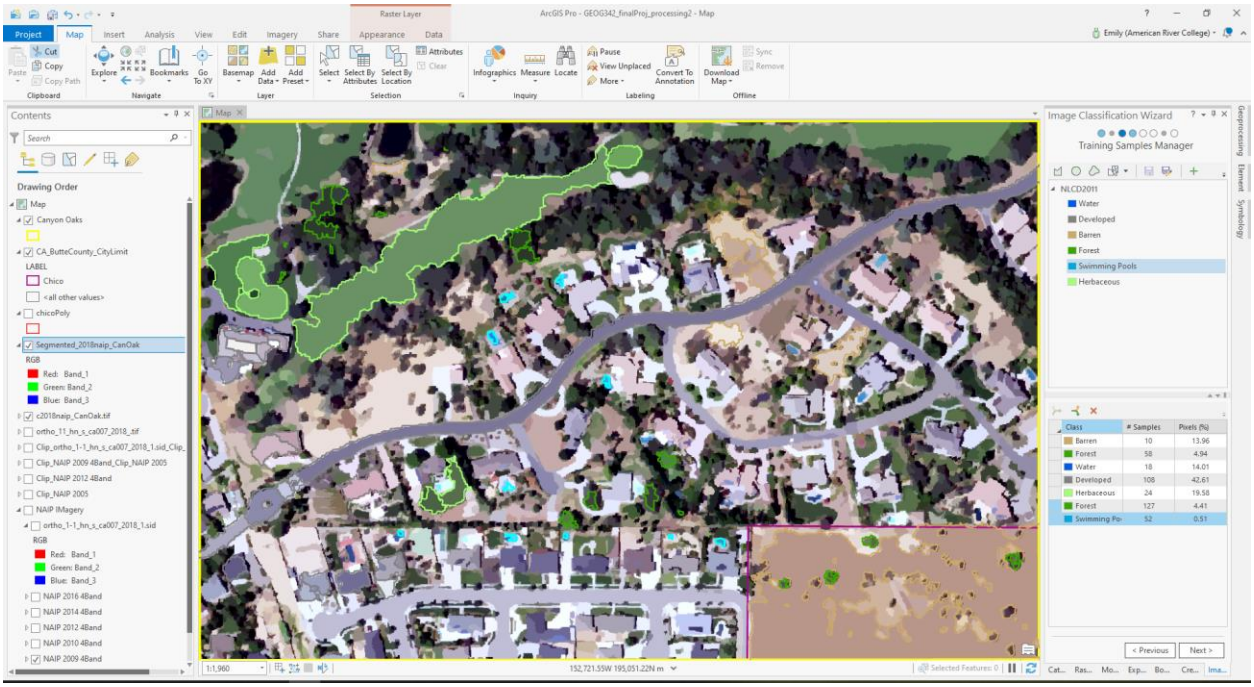


Figure 38. A zoom of the sampling from the previous figure.

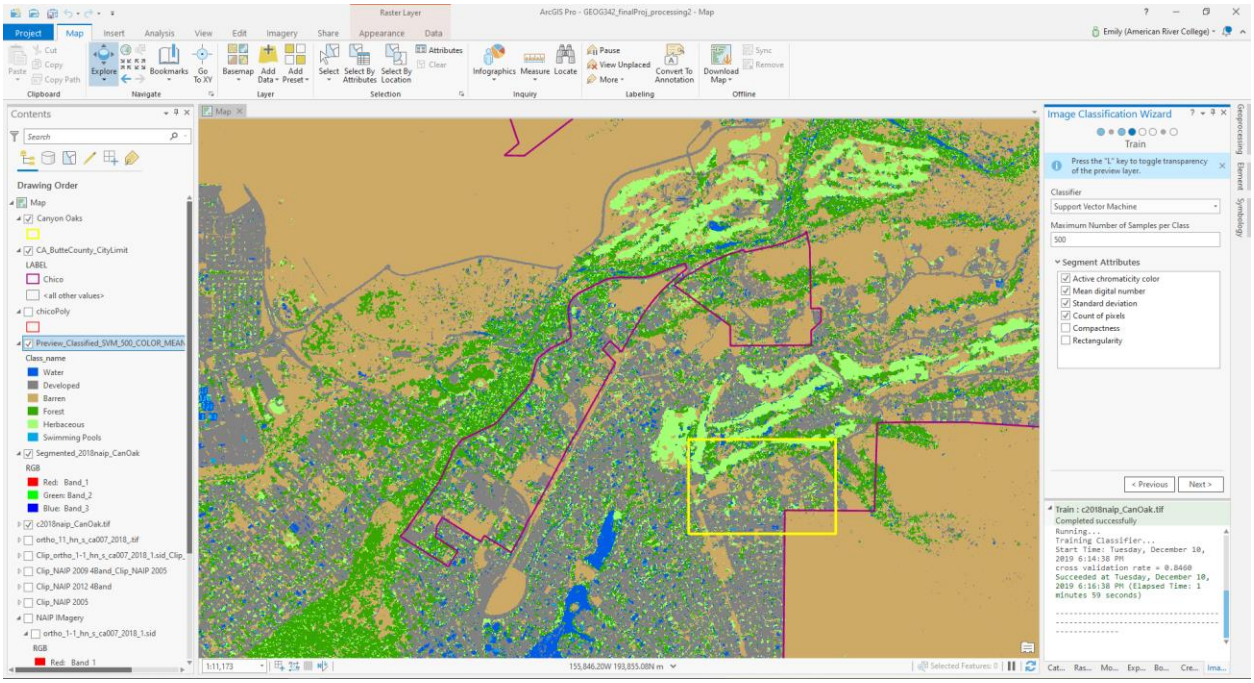


Figure 39. This set seems to work out pretty well, so I decided to move forward and classify. The moment of truth… would it work?

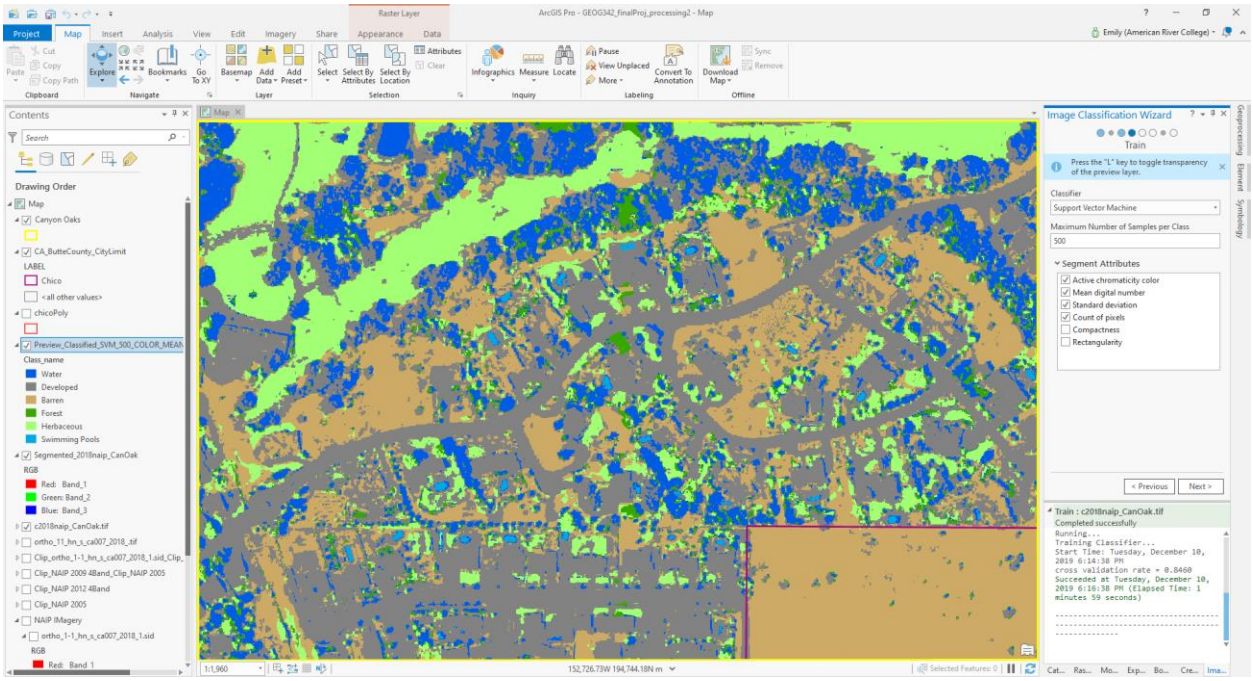


Figure 40. And another zoom of the Canyon Oaks area, for comparison...

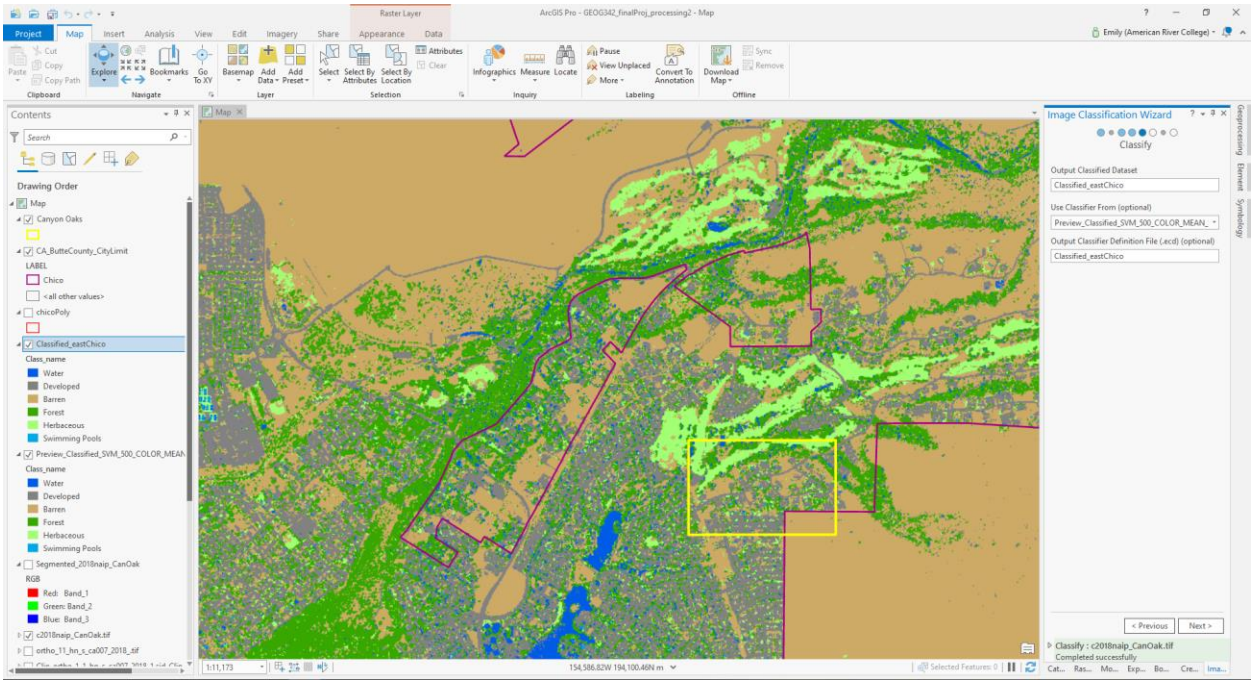


Figure 41. It worked!!!! The classification step actually ran this time.

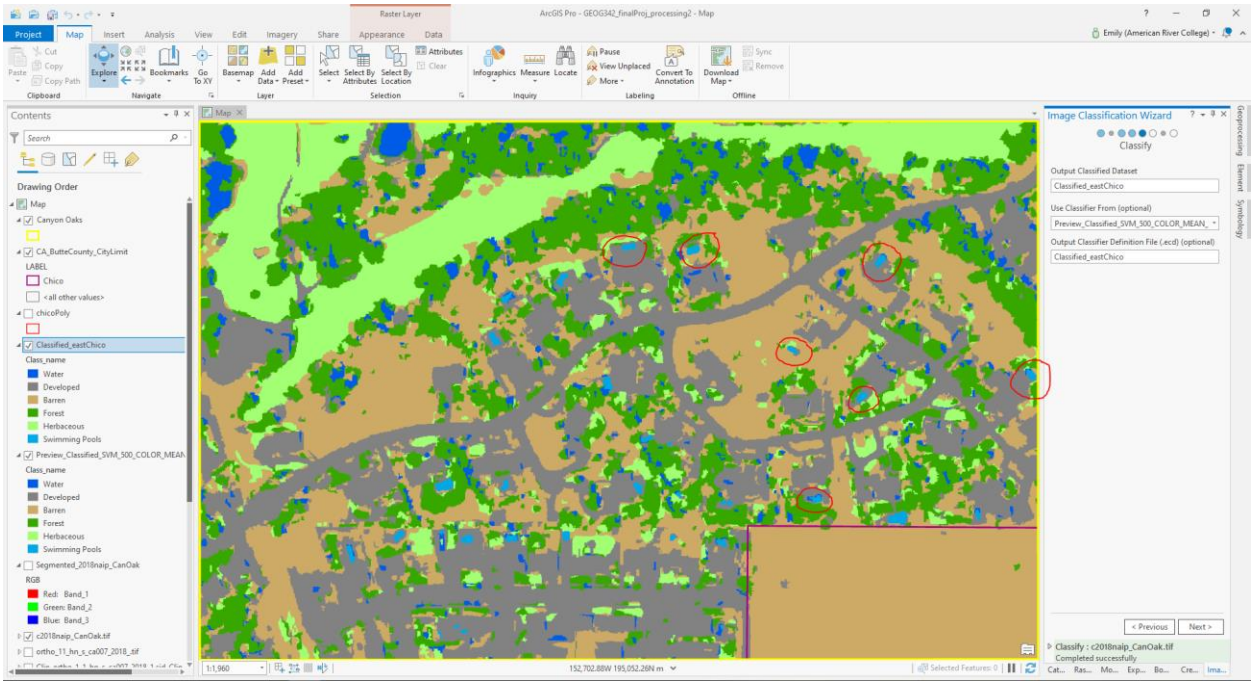


Figure 42. AND it picked up pools in the Canyon Oaks area that I didn't select! Though, it seems to think that Chico is a floating city. I'm more than happy with the result though.

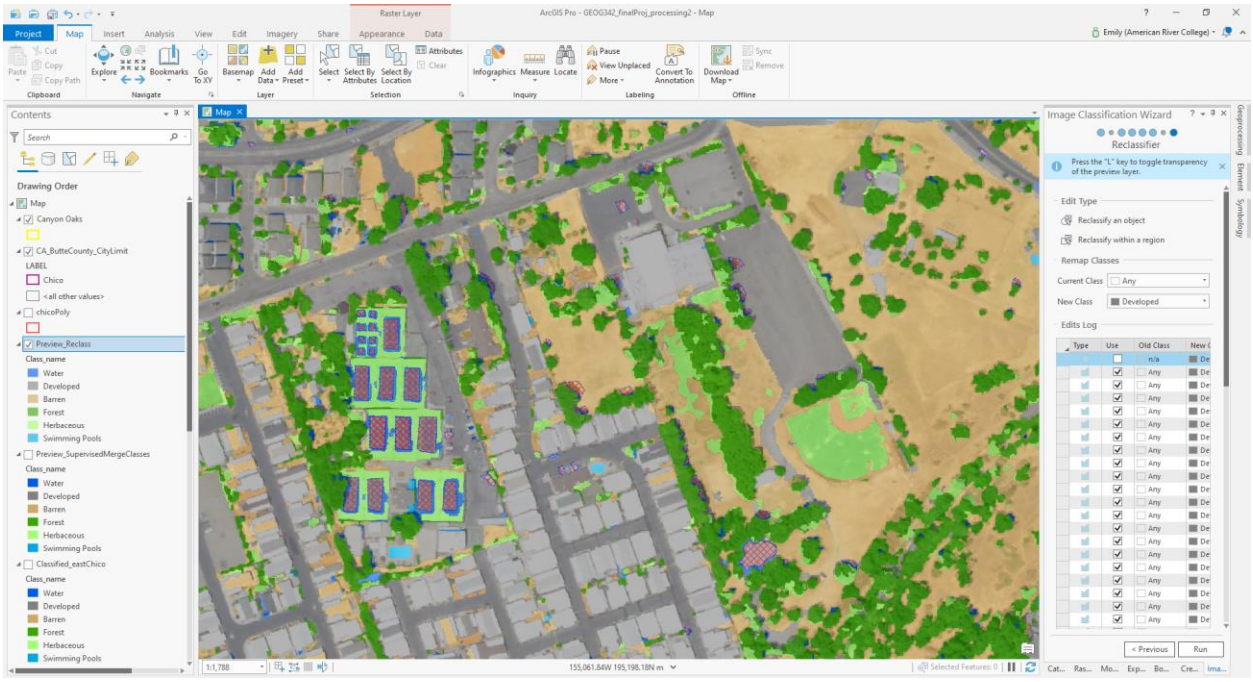


Figure 43. This image is a sample area where I reclassified "water" to "barren" or "developed" as needed. I gave the classified image some transparency so that it was easier to see what was below.

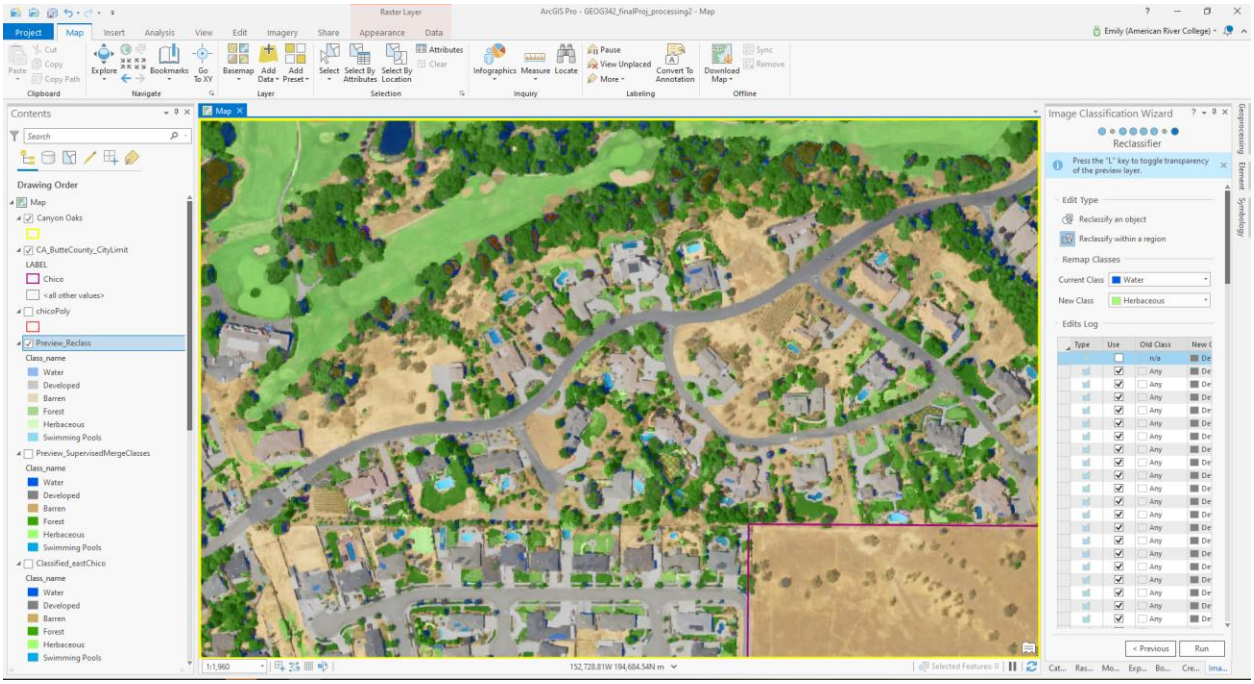


Figure 44. I only did a partial reclass on the Canyon Oak area, just to see what it would look like.

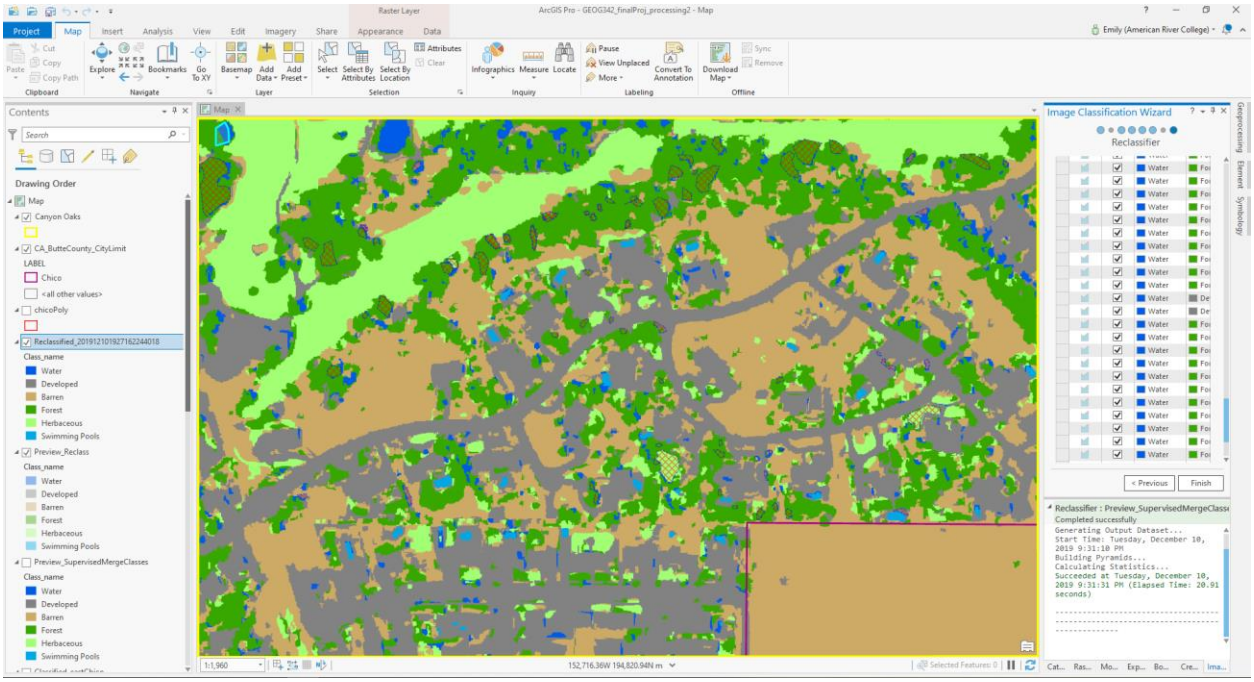


Figure 45. I ran the reclassifer and It did something!!!!

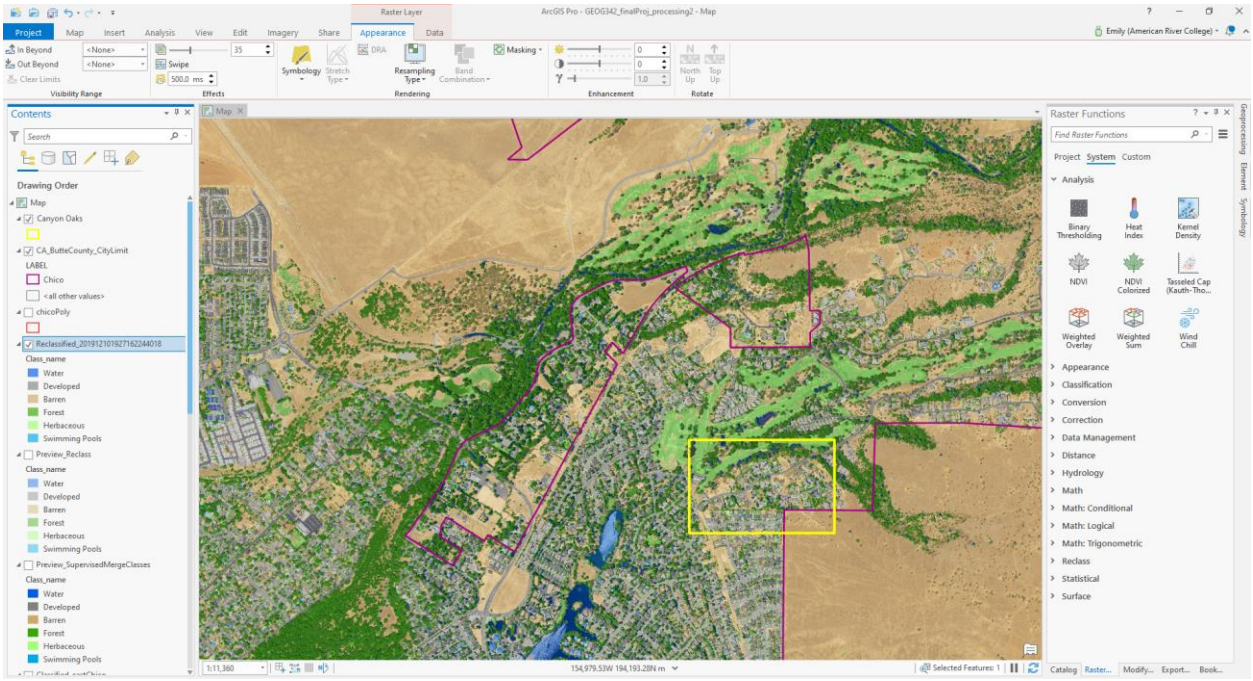


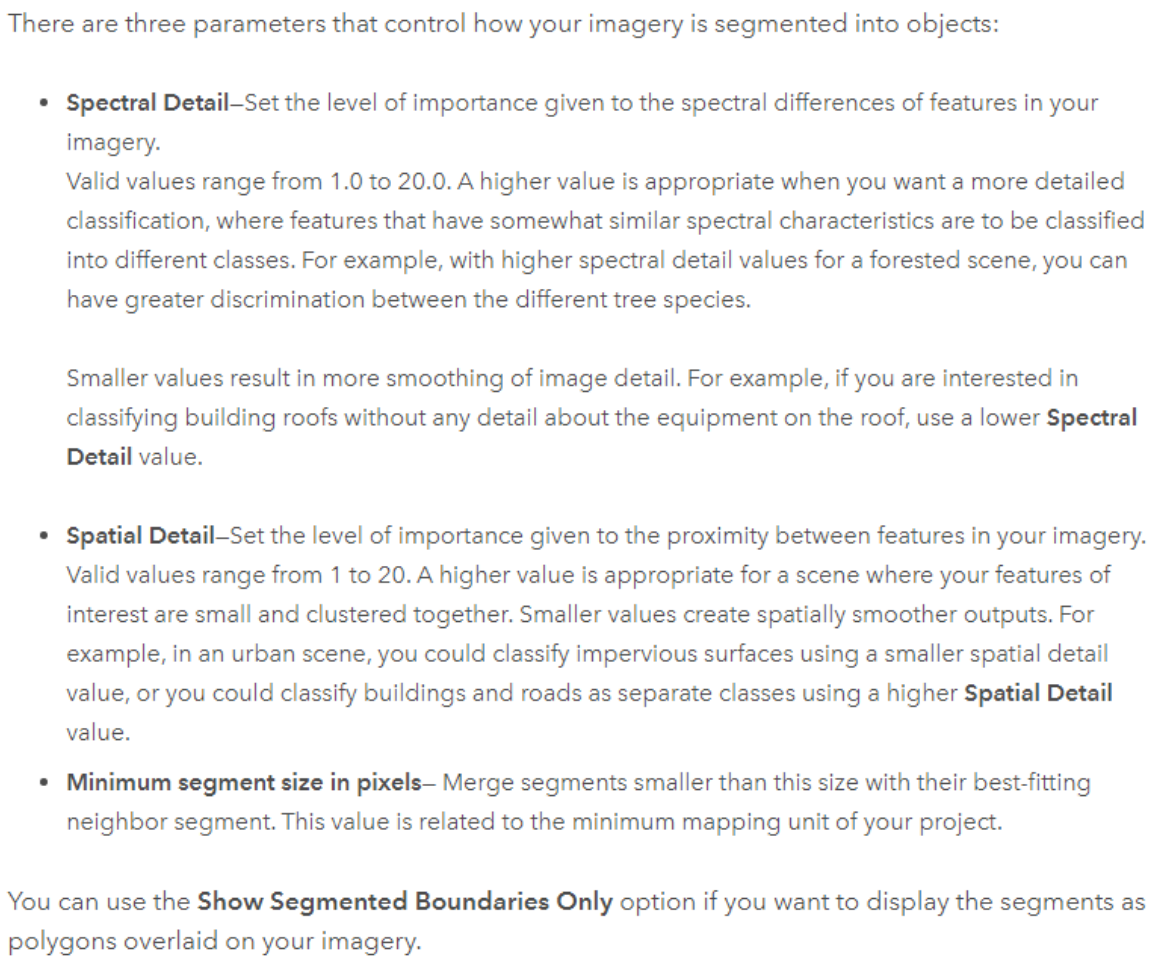
Figure 46. The 80th time is a charm! The final reclassified imaged with transparency, overlaid on the 2018 NAIP imagery.

Discussion

When initially tackling this project, I went full-in and set up a map with my desired base layers from the state webhosted layers, and then just started running a test using the defaults for the image segmentation and classification wizard tools from my flash drive. This process just led me to heartbreak and frustration, and a lot of "wasted" time, in terms of me starting the analysis, performing the training, the computer running for a full day, and then crashing when I attempt to make modifications. Trees were breaking down into individual segments when I was trying to run the classification tool, and I was collecting hundreds of samples for the training process. When I would try running the classification my computer would freeze up, and then I'd have to start all over with the image classification process.

I tried several things to help with the processing of my image, including basic steps like disabling sleep mode to prevent my computer from going into sleep mode while an image was processing, and saving the project to my hard drive as had been recommended with several of the exercises in class. However, this was not enough. Initially, I didn't realize that doing this processing directly from the webhosted layer wouldn't work. *Then* I found out that my 2018 NAIP was a .sid file, which, unbeknownst to me, was not a workable format for this project. Once I got on track and got the right kind of file (by exporting the clipped image to a .tif), and I removed all of the extra classes, this worked out pretty well. It would have been nice to have more time to work with this and try to eliminate more of the shadows, and to have the opportunity to do a second image to compare it to, but despite all that, it was DEFINITELY a learning experience. Though I might have learned the hard way, I've learned the mistakes not to make when doing this, and it's a lesson I'll not likely forget.

I was also struggling in the early stages to figure out how to best segment my image, since I was looking for such detailed features. So I went online and Googled some help, and came up with the following website: https://pro.arcgis.com/en/pro-app/help/analysis/image-analyst/the-image-classification-wizard.htm. What was most helpful with my classification in terms of the choices I should be making, was this information:



There are three parameters that control how your imagery is segmented into objects:

- **Spectral Detail**–Set the level of importance given to the spectral differences of features in your imagery.
  Valid values range from 1.0 to 20.0. A higher value is appropriate when you want a more detailed classification, where features that have somewhat similar spectral characteristics are to be classified into different classes. For example, with higher spectral detail values for a forested scene, you can have greater discrimination between the different tree species.

  Smaller values result in more smoothing of image detail. For example, if you are interested in classifying building roofs without any detail about the equipment on the roof, use a lower **Spectral Detail** value.

- **Spatial Detail**–Set the level of importance given to the proximity between features in your imagery. Valid values range from 1 to 20. A higher value is appropriate for a scene where your features of interest are small and clustered together. Smaller values create spatially smoother outputs. For example, in an urban scene, you could classify impervious surfaces using a smaller spatial detail value, or you could classify buildings and roads as separate classes using a higher **Spatial Detail** value.
- **Minimum segment size in pixels**– Merge segments smaller than this size with their best-fitting neighbor segment. This value is related to the minimum mapping unit of your project.

You can use the **Show Segmented Boundaries Only** option if you want to display the segments as polygons overlaid on your imagery.

Figure 47. Screenshot from an ESRI help site on the "Wizard Classification" tool, with guidance on the parameter options when performing the segmentation portion of the analysis.

Armed with new insight, I decided to tackle my image again. Using the preview option as an aid, I had the ability to see what the output segmented image would look like, and with the "Show Segment Boundary Options," I was able to see the underlying features while doing so to get an idea of how it was going to look once I started selecting training samples. This was really helpful to me in terms of getting a handle on what the different settings could do for my image, and how I could manage to get the level of detail I needed to capture small items like swimming pools, while reducing the number of segments that came up in the orchards or other areas with trees. What was really helpful was that as you made adjustments, the layer would update so you could see what you were losing – or gaining – with each modification. Below are several of the options I tried out before settling on my final segmentation scheme. Figure 48 shows the default settings that come up with the "Segmentation" tool. Figure 49 and Figure 50 show several of the modification I made when trying to determine how to segment my final image, while Figure 51 shows the final settings I chose. All images have the option checked for showing the segment boundaries only so that the underlying imagery can easily be viewed.

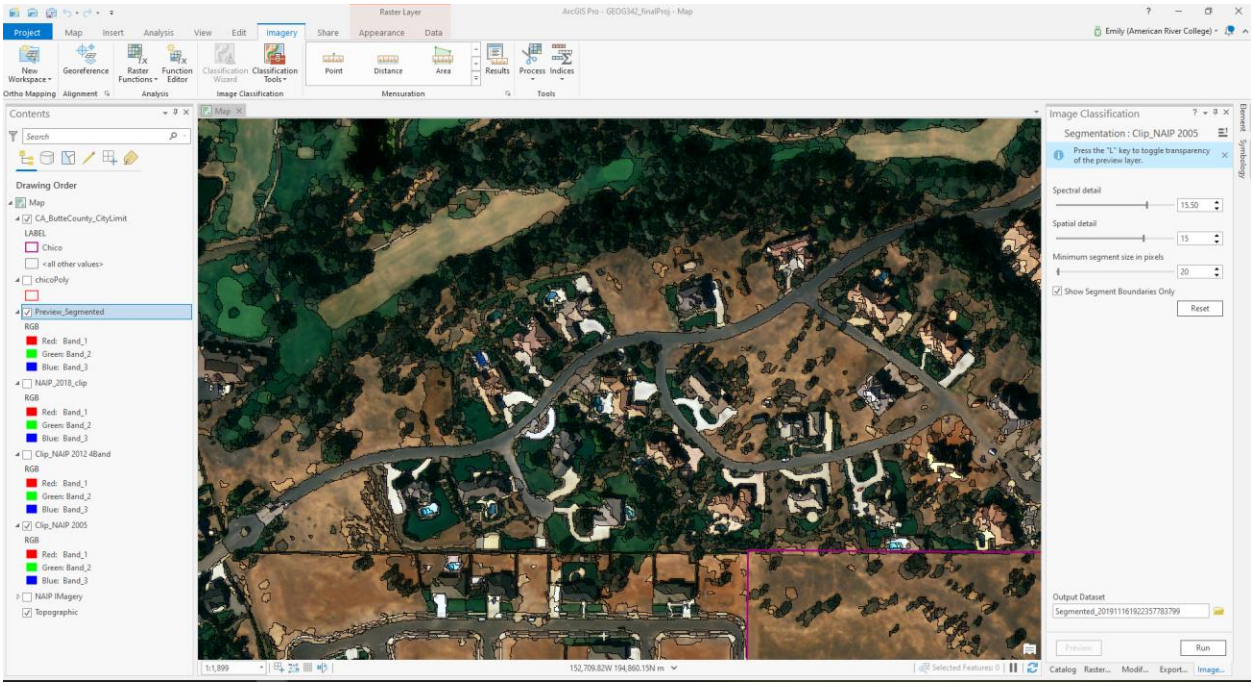


Figure 48. Screen shot of the 2005 image in ArcPro with the default settings that automatically prepopulate. This option produced an overly segmented view of the area and was more detailed than desired for the image classification.

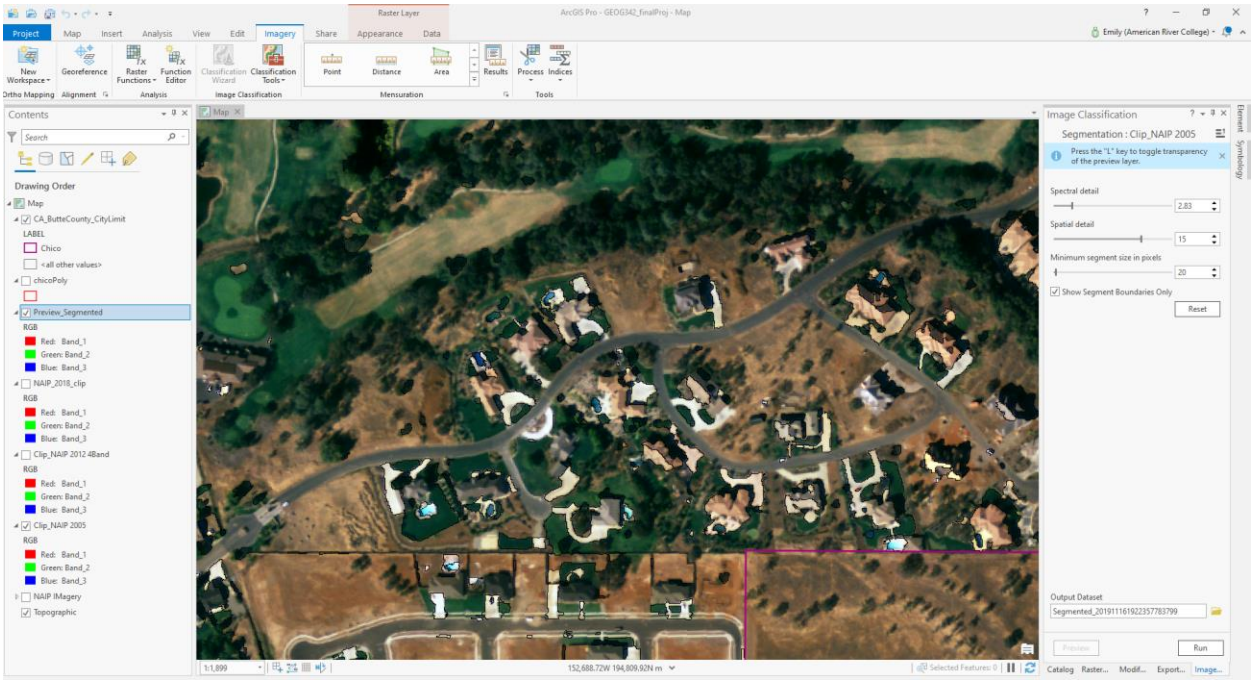


Figure 49. Screen shot of the 2005 image in ArcPro with settings for low spectral detail, higher spatial detail, and a minimum pixel size of 20. This option produced an overly generalized view of the area and was not detailed enough to perform the classification.

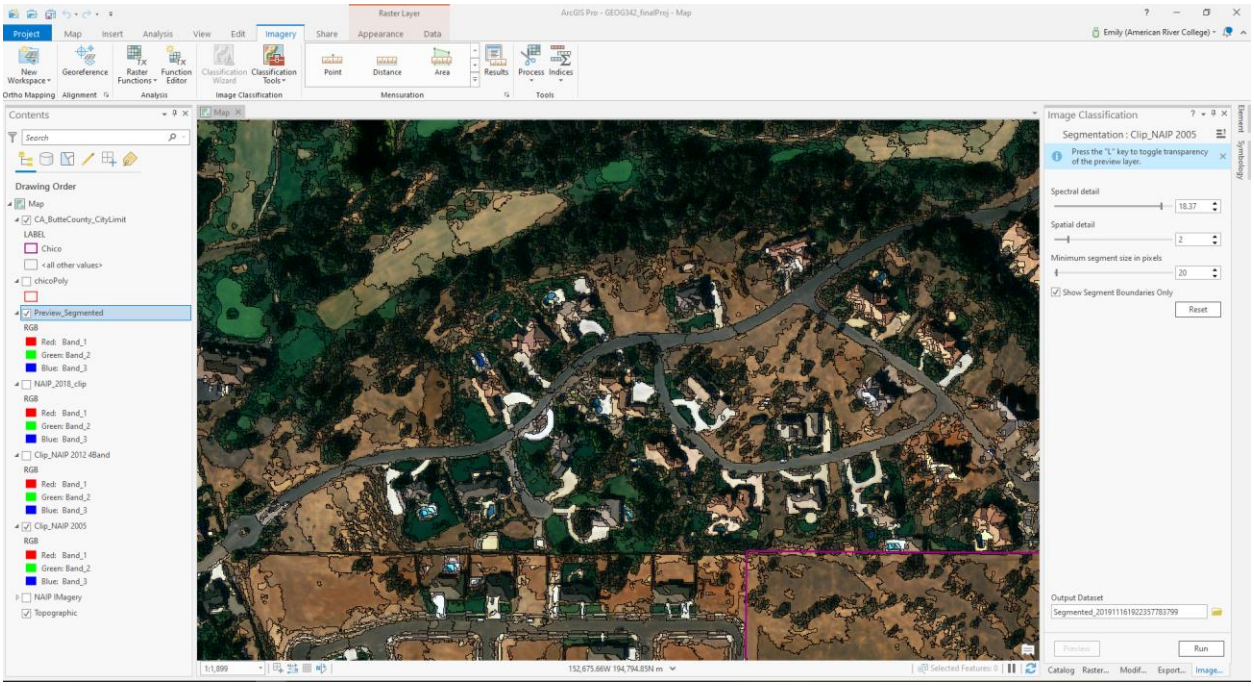


Figure 50. Screen shot of the 2005 image in ArcPro with settings for high spectral detail, lower spatial detail, and a minimum pixel size of 20. This option produced too detailed of a view of the area and was more than what was needed for this classification

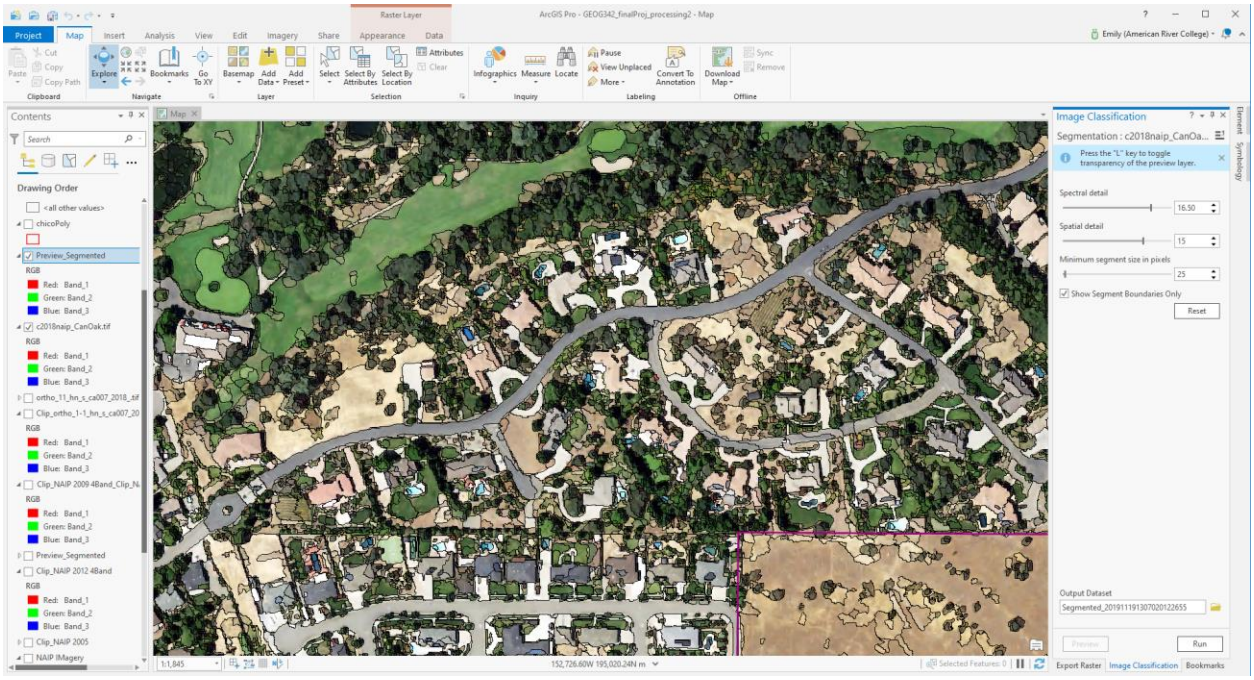


Figure 51. The compromise between the level of detail I wanted, and what I was willing to sacrifice for the 2005 image processing.

Other hard lessons learned while doing this project: orchards are their own nightmare, in that yes, they are ag lands, but yes, they are trees. When I was still trying to work with the entire City of Chico, I was struggling with all of the orchards that are on the edge of town because I wanted to classify them as their own thing… my theory being that maybe spatial distribution patterns and the reflectance values from the trees would somehow be different from those that were in the parks or people's yards. This was definitely not the case, and I learned that if I were to continue with this project and my

attempts at classifying the entire Chico area, they were going to have to go into my forested categories. Since I ended up dropping these areas it just became one less thing for me to worry about in the end.

Unfortunately, I never did have the time to run through this process again to do a second image for a comparison of the development of Chico over the past few years. I was really excited to learn that it was possible to pick up something as small as a swimming pool, and I think with the right classification scheme, it'd be possible to get even more out of the classes I had, but with modifications. For example, the rooftops in the area were so varied: brick reds, greys, and black; sloped and flat; chimneys, solar panels, all of the above. I almost wonder if one were to do a rooftop classification, if it would be better to have MORE categories, so that they could be separated by at least the color scheme, and have less variation going all into the same class. The trees might have eventually been able to be sorted out as well, but I think that it might be best to pick ONE thing to try and draw from these classifications, and not make them overly complicated.

References

Butte County Department of Water and Resource Conservation. (2016, June). *Butte County Water Inventory Analysis*. Retrieved November 17, 2019, from https://www.buttecounty.net/wrcdocs/Reports/I%26A/Report/4Climate%26Hydrology.pdf

ESRI. (n.d.). *The Image Classification Wizard*. Retrieved November 16, 2019, from https://pro.arcgis.com/en/pro-app/help/analysis/image-analyst/the-image-classification-wizard